Equilibrium statistical mechanics and energy partition for the shallow water model

Antoine Renaud, Antoine Venaille, Freddy Bouchet

Journal of Statistical Physics manuscript No.
(will be inserted by the editor)

A. Renaud, A. Venaille, F. Bouchet

Equilibrium statistical mechanics and energy partition for the shallow water model

Received: / Accepted:

Abstract The aim of this paper is to use large deviation theory in order to compute the entropy of macrostates for the microcanonical measure of the shallow water system. The main prediction of this full statistical mechanics computation is the energy partition between a large scale vortical flow and small scale fluctuations related to inertia-gravity waves. We introduce for that purpose a semi-Lagrangian discrete model of the continuous shallow water system, and compute the corresponding statistical equilibria. We argue that microcanonical equilibrium states of the discrete model in the continuous limit are equilibrium states of the actual shallow water system.

We show that the presence of small scale fluctuations selects a subclass of equilibria among the states that were previously computed by phenomenological approaches that were neglecting such fluctuations. In the limit of weak height fluctuations, the equilibrium state can be interpreted as two subsystems in thermal contact: one subsystem corresponds to the large scale vortical flow, the other subsystem corresponds to small scale height and velocity fluctuations. It is shown that either a non-zero circulation or rotation and bottom topography are required to sustain a non-zero large scale flow at equilibrium.

Explicit computation of the equilibria and their energy partition is presented in the quasi-geostrophic limit for the energy-enstrophy ensemble. The possible role of small scale dissipation and shocks is discussed. A geophysical application to the Zapiola anticyclone is presented.

Keywords Equilibrium statistical mechanics · shallow water model · large deviations · turbulence · inertia-gravity waves · geostrophic flows

Laboratoire de Physique de l'École Normale Supérieure de Lyon, Université de Lyon, CNRS, 46 Allée d'Italie, F-69364 Lyon CEDEX 07, France. antoine.renaud@ens-lyon.fr, antoine.venaille@ens-lyon.fr, freddy.bouchet@ens-lyon.fr

Contents

1 Introduction . . . 3
2 Shallow water model . . . 6
  2.1 Dynamics . . . 6
  2.2 Stationary states . . . 8
  2.3 Conserved quantities . . . 9
3 Equilibrium statistical mechanics of a discrete shallow water model . . . 10
  3.1 Liouville theorem . . . 11
  3.2 A discrete model . . . 12
    3.2.1 Definition of the ensemble of microscopic configurations . . . 13
    3.2.2 Coarse-graining . . . 14
    3.2.3 Definition of a velocity field on the grid . . . 14
    3.2.4 Definition of the microcanonical ensemble for the discrete model. . . . 15
  3.3 Macrostates and their entropy . . . 18
    3.3.1 Definition of the empirical density field . . . 18
    3.3.2 Definition of the macrostates . . . 18
    3.3.3 Macroscopic observables and empirical density . . . 19
    3.3.4 Asymptotic behavior of the macrostates volume and derivation of its entropy . . . 19
  3.4 Continuous limit . . . 21
    3.4.1 Expressions of the macrostate entropy, energy and potential vorticity moments . . . 21
    3.4.2 Microcanonical variational problem for the probability density field . . . 22
4 General properties of the equilibria and simplification of the theory in limiting cases. . . . 23
  4.1 Equilibrium states are stationary states with positive temperatures . . . 24
    4.1.1 Properties of the critical points . . . 24
    4.1.2 Equation for the large scale flow . . . 25
    4.1.3 Comparison with previous results . . . 26
  4.2 Equipartition and decoupling in the limit of weak local height fluctuations . . . 26
    4.2.1 The height distribution . . . 26
    4.2.2 Equipartition and the limit of weak small scale energy . . . 27
    4.2.3 Decoupling between the large scale flow and the fluctuations . . . 27
    4.2.4 Either a non-zero circulation or a non-zero bottom topography is required to sustain a large scale flow at equilibrium . . . 29
  4.3 The quasi-geostrophic limit . . . 29
    4.3.1 Geostrophic balance . . . 29
    4.3.2 Quasi-geostrophic dynamics . . . 30
    4.3.3 Quasi-geostrophic constraints and variational problem . . . 31
    4.3.4 Maximum energy states and consistency of the quasi-geostrophic approximation. . . . 32
5 Explicit computation of phase diagrams and discussion . . . 33
  5.1 Energy-enstrophy equilibria for the quasi-geostrophic model. . . . 33
  5.2 The effect of energy dissipation and enstrophy dissipation . . . 35
  5.3 Flow structure of the equilibrium states: application to the Zapiola Anticyclone. . . . 37
6 Conclusion . . . 37
A Invariant measure and formal Liouville theorem . . . 40
  A.1 Formal Liouville theorem for the triplet of fields (h,hu,hv) . . . 40
  A.2 Change of variables from (h,hu,hv) to (h,q,μ) . . . 41
B Relevance of the constraints for the discrete model . . . 41
  B.1 Areal coarse-graining for continuous fields . . . 41
  B.2 Potential vorticity moments . . . 42
  B.3 Energy . . . 42
C Critical points of the mean-flow variational problem . . . 43
  C.1 Intermediate variational problem . . . 44
  C.2 Computation of the critical points . . . 45
D Global maximizers of the entropy of the large scale flow . . . 47
E Comparison with a Eulerian discrete model . . . 48
F Energy-Enstrophy ensemble . . . 49
  F.1 Computation of the critical points . . . 49
  F.2 Construction of the phase diagram . . . 50
  F.3 Limit cases for the energy partition . . . 50

1 Introduction

Geophysical turbulent flows have the propensity to self-organize into large scale coherent structures such as cyclones, anticyclones and jets. These coherent structures are long lived, but can also loose energy, for instance through the radiation of waves that eventually break into small scale structures. The aim of this paper is to understand the energy partition into large scale structures and small scale fluctuations in the framework of freely evolving shallow water dynamics, using statistical mechanics arguments. Indeed, geophysical turbulent flows involve a huge number of degrees of freedom coupled through non-linear interactions, which strongly motivates a statistical mechanics approach. This approach allows to reduce the study of self-organization and energy partition down to a few parameters, such as the total energy of the flow and its total enstrophy.

In the case of the three dimensional Euler equations, equilibrium statistical mechanics predicts that all the energy is lost into small scales, consistently with the classical picture of a small scale energy transfer. By contrast, two-dimensional flows are characterized by a large scale energy transfer, and equilibrium tools are appropriate to describe the large scale structure resulting from self-organization at the domain scale, in the absence of forcing and dissipation. The idea goes back to Onsager [28], and has been mostly developed during the nineties after the work of Miller-Robert-Sommeria [23,33], see also Refs. [17,10,4] and references therein. Importantly, the theory predicts that the contribution of small scale fluctuations to the total energy are negligible in the two-dimensional case. Equilibrium statistical mechanics of two-dimensional and quasi-geostrophic flows is now fairly well understood. It has been applied to various problems in geophysical context such as the description of Jovian vortices [44,3], oceanic rings and jets [50,55], equilibria on a sphere [15], and to describe the vertical energy partition in continuously stratified quasi-geostrophic flows [19,52,48].

Due to the combined effect of stable stratification, thin aspect ratio and rotation, geophysical flows are very different from classical three-dimensional turbulence. However, such flows are not purely two-dimensional. Here we consider the shallow water equations, which is an intermediate model between three-dimensional and two-dimensional turbulence. This model describes the dynamics of a thin layer a fluid with homogeneous density. On the one hand, shallow water equations admit conservation laws similar to two-dimensional Euler equations, that lead to self-organization of the energy at large scale in the Euler case. On the other hand, shallow water dynamics support the presence of inertia-gravity waves that are absent from purely two-dimensional turbulence. A small scale energy transfer may exist due to the existence of these inertia-gravity waves in the shallow water system. The quasi-geostrophic model is recovered as a limit case of the shallow water model, when the Rossby parameter (comparing inertial terms to Coriolis forces) is small. A small Rossby number corresponds to a strong rotation limit. It is then natural to ask whether previously computed statistical equilibrium states of the quasi-geostrophic models remain equilibrium states of the shallow water model. More generally, given a certain amount of energy in an unforced, undissipated geophysical flow, will the flow self-organize into a large scale coherent structure, just as in two dimensional turbulence? Or will the energy be transferred towards the small scales, just as in three dimensional homogeneous turbulence? The aim of this paper is to answer these questions by computing statistical equilibrium states of the inviscid shallow water model.

The first step before computing equilibrium states is to identify a suitable phase space to describe microscopic configurations of the system. The phase space variables must satisfy a Liouville theorem, which ensures that the flow in phase space is divergence-less. Consequently, a uniform measure on a constant energy-Casimirs shell of phase space is invariant (microcanonical measure). The second step is to describe the system at a macroscopic level. The macrostates will be the sets of microstates sharing the same macroscopic behavior. The third step is to find the most probable macrostate, and to show that almost all the microstates correspond to this macrostate for given values of the constraints. While these three steps may be straightforward for differential equations with a finite number of degrees of freedom, for continuous systems described by partial differential equations, these three steps require the introduction of discrete approximations of the continuous field and of the invariant measure, and to study the continuous field limit of these discrete approximations. This point will be further discussed in the following.

This program has been achieved in the 90s for the two-dimensional Euler equations. Indeed, a Liouville theorem is satisfied by the vorticity field, which describes therefore a microscopic configuration of the system. A macrostate can be defined as a probability field describing the distribution of vorticity levels at a given point, either through a coarse graining procedure [22, 23] or directly by the introduction of Young measures [30, 31, 33].

For the shallow water dynamics, they are further issues that need to be overcome. The existence of a Liouville theorem for the shallow water flow was found by Warn [53] by describing the flow configurations on a basis given by the eigenmodes of the linearized dynamics. However, the constraints of the problem given by dynamical invariants are not easily expressed in terms of the variables satisfying this Liouville theorem (except in the weak flow limit discussed by Warn [53]). This difficulty has been overcome by Weichman and Petrich [56] who considered first a Lagrangian representation of the flow, and then used a formal change of variable to describe the flow configurations with Eulerian variables convenient to express the constraints of the problem. Using a different method that does not require a Lagrangian representation of the fluid, we will show the existence of a formal Liouville theorem.

A second difficulty concerns the choice of a relevant discrete approximation that allows to keep as much as possible geometric conservations laws of the continuous dynamics. Those geometric conservation laws include the Liouville property, the Lagrangian conservation laws (i.e. the conservation of the volume carried by each fluid particle), and global dynamical invariants. Unfortunately, we are not aware of a discrete model that does not break at least one of those geometric conservation laws. However, we will argue that there is no logical need for the discrete model to satisfy all the conservation laws of the continuous dynamics in order to guess the correct microcanonical measure of the continuous system by considering the limit of a large number of degrees of freedom.

A third difficulty is that local small scale fluctuations of the fields may have a substantial contribution to the total energy. This contrasts with 2d Euler dynamics, for which small scale fluctuations of the vorticity field have a vanishingly small contribution to the total energy in the continuous limit. In the shallow water case, it is not a priori obvious that the contribution of these small scale fluctuations to the total energy can be expressed in terms of the macroscopic probability density field.

In order to overcome the second and the third difficulty, we introduce in this paper a semi-Lagrangian discretization : one the one hand, we consider fluid particles of equal volume, which allows to keep track of the Lagrangian conservation laws of the dynamics. On the other hand, the particle positions are restricted to a uniform horizontal grid, and each grid node may contain many particles, which allows for an Eulerian representation of the macrostates and to keep track of global conservation laws, while taking into account the presence of small scale fluctuations contributing to the total energy. We then derive the statistical mechanics theory for this discrete representation of the shallow water model, using large deviation theory.

We will argue that in the limit of a large number of degrees of freedom, the equilibrium states of the semi-Lagrangian discrete model are the equilibrium states of the actual shallow water system, and we will also show that we recover with this model results already obtained in several limiting cases. While the statistical mechanics treatment of our discrete model is rigorous, there is some arbitrariness in our definition of the discrete model, which is not fully satisfactory.

We stress that our approach to guess the invariant measure of the continuous shallow-water system is heuristic: there is to our knowledge no simple way to define mathematically the microcanonical measure or the Gibbs measure of an Hamiltonian infinite dimensional system. As far as we know the only example of a rigorous work for defining invariant measures in the class of deterministic partial differential equations of interest, is a work by [5] on the periodic nonlinear Schrodinger equation. For the 2d Euler equations, the microcanonical measure seems clear from a physical point of view because different discretizations, for instance either Eulerian ones [22, 32] or Lagrangian ones (see [8], lead to a consistent picture in the thermodynamics limit, but even in that case no clear mathematical construction of the invariant measure exists.

There exist only a very limited number of results on statistical equilibrium states of the shallow water system. Warn [53] studied the equilibrium states in a weak flow limit in the energy-enstrophy ensemble. He showed that in the absence of lateral boundaries and in the absence of bottom topography, all the initial energy of a shallow water flow is transferred towards small scales. Here we relax the hypothesis of a weak flow, generalizing the conclusions of Warn [53] for any flow, and we discuss the effect of bottom topography. We show that when there is a non zero bottom topography and when the flow is rotating, there is a large scale flow associated with the equilibrium states.

Refs. [20,7] did compute statistical equilibrium states of “balanced” shallow water flows, by assuming that all the energy remain in the large scale flow. The equilibria described by Merryfield et al [20] were obtained in the framework of the energy-enstrophy theory, neglecting any other potential vorticity moments than the potential enstrophy. Similar states were described as minima of the potential enstrophy for the macroscopic potential vorticity field by Sanson [37].

Chavanis and Sommeria [7] proposed a generalization of the 2D-Euler variational problem given by the Miller-Robert-Sommeria theory to the shallow water case. Their main result is a relationship between the large scale streamfunction and height field. This result was very interesting and inspiring to us. However, Chavanis and Sommeria did not derive their variational problem from statistical mechanics arguments but proceeded through analogy. They were moreover neglecting height fluctuations. This does not allow for energy partition between vortical flow and fluctuations, and moreover leads to inconsistencies for some range of parameters, as the Chavanis-Sommeria constrained entropy can be shown to have no maxima for negative temperatures (the negative temperature critical points of this variational problem are saddles rather than maxima). As a consequence a range of possible energy values can not be achieved in this phenomenological framework. Nevertheless, with an approach based on statistical mechanics, we will confirm, in this study, the form of the variational problem proposed by Chavanis and Sommeria [7] for describing part of the field, for a restricted range of parameters. We stress however, that generically part of the energy will be carried by the fluctuations, and thus that the Chavanis–Sommeria variational problem should be considered with an energy value which is not the total one. Determining which part of the energy should be taken into account requires a full statistical mechanics treatment, taking into account height fluctuations. We also note that for some other ranges of parameters, all the energy will be carried by the fluctuations, and thus the Chavanis-Sommeria variational problem then does not make sense.

In the preparation of this work, we have also been inspired by the work of Weichman and Petrich [56]. In this work, the authors computed a class of statistical equilibrium states of the shallow water system, starting from a grand canonical distribution. The main result of their work is the same equation describing the relation between the large scale stream function and height field, as the one previously obtained by Chavanis and Sommeria [7]. The Weichman–Petrich approach seemed more promising than the Chavanis–Sommeria one as it was based on a statistical mechanics treatment. However it also fails to predict which part of the energy goes into fluctuations, and to recognize the range of parameters for which the mean field equation for the largest scale is relevant. The reason is that, while natural in condensed matter physics, the hypothesis of a Gibbs or a grand canonical distribution, is questionable for an inertial flow equation which is in contact neither with an energy bath, nor with a potential vorticity bath. Related to this issue, Weichman and Petrich had to assume an ad hoc scaling for the thermodynamical parameters, in order to obtain statistical ensembles where entropy actually balances conservation laws. It has been recognized for a long time that this kind of theoretical difficulties are related to the Rayleigh-Jeans paradox. The proper way to address these issues, in some fluid models like for instance the two-dimensional Euler equation, is to start from the microcanonical measure rather than from a canonical or grand canonical one [4]. We will overcome all these problems in this work by solving the microcanonical problem, which may seem more difficult, but which can be handled using large deviation theory. It also leads to a more precise description of macrostate probabilities.

In this paper, starting from the microcanonical measure, we propose a complete computation of the macrostate entropy for the shallow water equations. In order to achieve this goal, we consider a large-deviation approach that allows to perform the statistical mechanics computation in an explicit and clear way, and that gives a precise description of most-probable macrostates. Moreover, we compute explicitly both the large scale flow and the small scales fluctuations of the equilibrium states. Our results are thus a complete statistical mechanics treatment of the shallow water equations.

We also connect our general results to some of the partial treatment discussed in the previous literature. For instance, we show that only one subclass of the states described by [20,56,7] are actual statistical equilibria of the shallow water model. More precisely, we will see that the large scale flow and the small scales fluctuations can be interpreted as two subsystems in thermal contact, and that the temperature of small scales fluctuations is necessarily positive. The large scale flow is therefore also characterized by a positive temperature.

The paper is organized as follows. The shallow water model and its properties are introduced in section 2. Equilibrium statistical mechanics of a discrete flow model arguably relevant to describe the continuous shallow water system is derived in section 3. Computation of the equilibrium states and a description of their main properties are presented in section 4. Energy partition between a large scale flow and small scale fluctuations is computed analytically in the quasi-geostrophic limit in section 5, which also includes a geophysical application to the Zapiola anticyclone. The main results are summarized and discussed in the conclusion.

2 Shallow water model

2.1 Dynamics

The shallow water equations describe the dynamics of a fluid layer with uniform mass density, in the limit where the layer depth is very small compared to the horizontal length scales of the flow [46,29]. In this limit the vertical momentum equation yields hydrostatic equilibrium, and the horizontal velocity field $(u,v) = \mathbf{u}(\mathbf{x},t)$ is depth independent, where $\mathbf{x} = (x,y)$ is any point of a two-dimensional simply connected domain $\mathscr{D}$. We consider the Coriolis force in the f-plane approximation, i.e. with a constant Coriolis parameter f and a rotation axis along the vertical direction. We denote $\eta(\mathbf{x},t)$ the vertical displacement of the upper interface and $h_b(\mathbf{x})$ the bottom topography (see Fig. 1). The origin of the z axis is chosen such that

$$\int_{\mathscr{D}} \mathrm{d}\mathbf{x}\, h_b(\mathbf{x}) = 0, \tag{1}$$

and the vertical displacement is defined such that

$$\int_{\mathscr{D}} \mathrm{d}\mathbf{x}\, \eta(\mathbf{x},t) = 0 \tag{2}$$

as well. We introduce the total depth

$$h = H - h_b + \eta(\mathbf{x},t), \tag{3}$$

where

$$H = \frac{1}{|\mathscr{D}|}\int_{\mathscr{D}} \mathrm{d}\mathbf{r}\, h(\mathbf{r}) \tag{4}$$

is the mean depth of the fluid, with $|\mathscr{D}|$ the area of the flow domain. The horizontal and vertical length units can always be chosen such that the domain area and the mean height H are equal to one ($|\mathscr{D}| = 1$, $H = 1$), and this choice will be made in the remainder of this paper.

The dynamics is given by the horizontal momentum equations

$$\begin{cases} \partial_t u + u\partial_x u + v\partial_y u - fv = -g\partial_x(h+h_b) \\ \partial_t v + u\partial_x v + v\partial_y v + fu = -g\partial_y(h+h_b) \end{cases} \tag{5}$$

and the mass continuity equation

$$\partial_t h + \nabla\cdot(h\mathbf{u}) = 0, \tag{6}$$

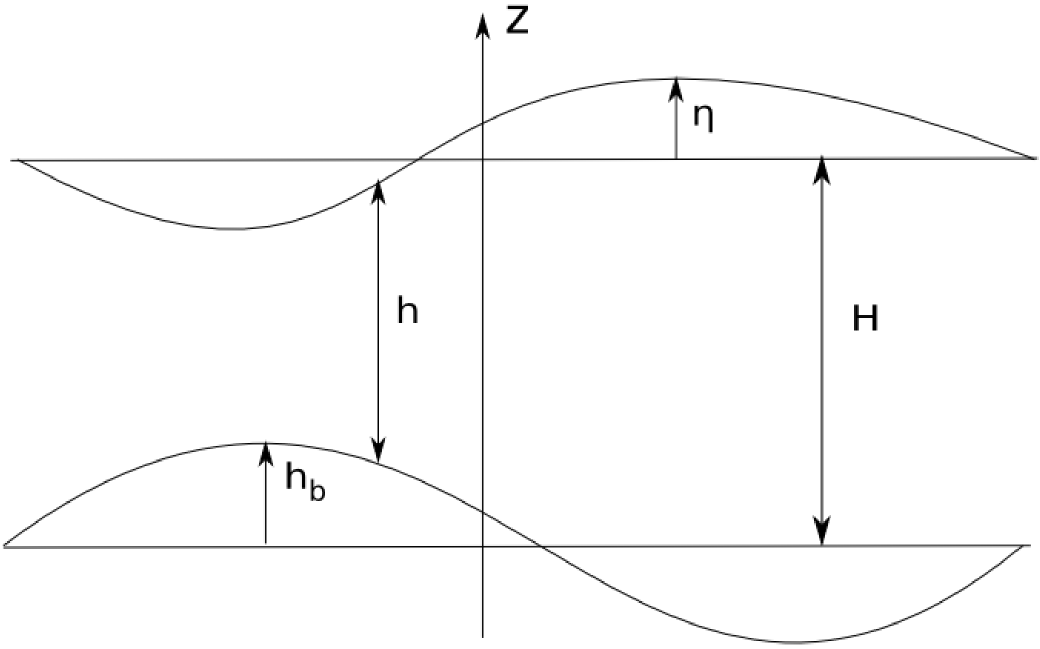

Fig. 1: Scheme of a vertical slice of fluid, η is the upper interface displacement, h_b is the bottom topography, h is the height field and H is the mean height of the fluid ($\langle h \rangle$)

where $\nabla = (\partial_x, \partial_y)$, with impermeability boundary conditions

$$\mathbf{u} \cdot \mathbf{n} = 0 \quad \text{on } \partial\mathscr{D}, \tag{7}$$

where $\mathbf{n}$ is the outward border-normal unit vector.

By introducing the Bernoulli function

$$B \equiv \frac{1}{2}\mathbf{u}^2 + g\,(h + h_b - 1)\,, \tag{8}$$

the relative vorticity field

$$\omega = \partial_x v - \partial_y u, \tag{9}$$

the divergence field

$$\zeta = \nabla \cdot \mathbf{u} = \partial_x u + \partial_y v, \tag{10}$$

and the potential vorticity field

$$q = \frac{\omega + f}{h}, \tag{11}$$

the momentum equations (5) can be recast into a conservative form

$$\begin{cases} \partial_t u - qhv = -\partial_x B \\ \partial_t v + qhu = -\partial_y B \end{cases} . \tag{12}$$

One can get from the momentum equations (12) a dynamical equations for the potential vorticity field (11) and for the divergence field (10), respectively:

$$\partial_t q + \mathbf{u} \cdot \nabla q = 0, \tag{13}$$

$$\partial_t \zeta - \nabla^{\perp} \cdot (qh\mathbf{u}) = -\Delta B, \tag{14}$$

where $\Delta = \nabla \cdot \nabla$ is the Laplacian operator and $\nabla^{\perp} = (-\partial_y, \partial_x)$. In order to show that the shallow water system described by $(h, \mathbf{u})$ is also fully described by the triplet (h, q, ζ), we introduce the velocity streamfunction $\psi(x, y, t)$ and the velocity potential $\phi(x, y, t)$ through the Helmholtz decomposition of the velocity[1]:

$$\mathbf{u} = \nabla^{\perp}\psi + \nabla\phi. \tag{15}$$

The boundary condition for the velocity given in Eq. (7) yields

$$\begin{cases} \mathbf{n} \cdot \nabla^{\perp}\psi = 0 \\ \mathbf{n} \cdot \nabla\phi = 0 \end{cases} \quad \text{on } \partial\mathscr{D}, \tag{16}$$

where $\mathbf{n}$ is the boundary normal vector. The relative vorticity (9) and the divergence (10) can be expressed in terms of the velocity streamfunction and velocity potential as

$$\omega = \Delta\psi, \quad \zeta = \Delta\phi \ . \tag{17}$$

To the fields ω and ζ correspond a unique field ψ and a field ϕ defined up to a constant with the boundary condition (16). Thus, to the fields $\omega = qh - f$ and ζ corresponds a unique velocity vector written formally as

$$\mathbf{u} = \nabla^{\perp}\Delta^{-1}(qh - f) + \nabla\Delta^{-1}\zeta. \tag{18}$$

It will also be useful to consider the field μ defined as

$$\mu = \Delta^{1/2}\phi, \tag{19}$$

where $\Delta^{1/2}$ is the linear operator whose eigenmodes are Laplacian eigenmodes in the domain $\mathscr{D}$ with the boundary condition $\mathbf{n} \cdot \nabla = 0$ on $\partial\mathscr{D}$, and whose eigenvalues are the negative square root of the modulus of Laplacian eigenvalues. The field μ can be interpreted as a measure of the amplitude of the potential contribution to the velocity field, given that $\int_{\mathscr{D}} \mu^2 = \int_{\mathscr{D}} (\nabla\phi)^2$. This field will also be referred to as the divergence field in the following. To the field μ correspond a unique field ϕ (up to a constant). Thus, the shallow water system is fully described by the triplet (h, q, μ), and the velocity vector can be written formally in terms of these fields as

$$\mathbf{u} = \nabla^{\perp}\Delta^{-1}[hq - f] + \nabla\Delta^{-1/2}[\mu]. \tag{20}$$

2.2 Stationary states

We investigate in this subsection the conditions for stationarity of the flow in the shallow water model following [7]. Using a Helmholtz decomposition, one can define the mass transport streamfunction $\Psi(x, y, t)$ and the mass transport potential $\Phi(x, y, t)$, not to be confused with ψ and ϕ defined in Eq. (15) and (16):

$$h\mathbf{u} = \nabla^{\perp}\Psi + \nabla\Phi. \tag{21}$$

The boundary condition for the velocity given in Eq. (7) yields

$$\begin{cases} \mathbf{n} \cdot \nabla^{\perp}\Psi = 0 \\ \mathbf{n} \cdot \nabla\Phi = 0 \end{cases} \quad \text{on } \partial\mathscr{D}. \tag{22}$$

[1] In the particular case of a bi-periodic domain, i.e. with periodic boundary conditions for the velocity $\mathbf{u}$, one would need to describe in addition the homogeneous part of the velocity field, which is both divergence-less and irrotational.

It will be useful in the remainder of the paper to express the vorticity field defined in Eq. (9) and the divergent field defined in Eq. (10) in terms of these transport streamfunction Ψ and transport potential Φ :

$$\omega = \nabla\cdot\left(\frac{1}{h}\nabla\Psi\right) + J\left(\frac{1}{h},\Phi\right), \tag{23}$$

$$\zeta = \nabla\cdot\left(\frac{\nabla\Phi}{h}\right) - J\left(\frac{1}{h},\Psi\right), \tag{24}$$

where

$$J(f,g) = \partial_x f \partial_y g - \partial_y f \partial_x g \tag{25}$$

is the Jacobian operator. The dynamics in Eqs. (6), (13) and (14) can also be written in terms of Ψ and Φ:

$$\partial_t h = -\Delta\Phi, \tag{26}$$

$$\partial_t q + \frac{1}{h}J(\Psi,q) + \frac{1}{h}\nabla\Phi\cdot\nabla q = 0, \tag{27}$$

$$\partial_t\zeta + J(\Phi,q) - q\Delta\Psi - \nabla q\cdot\nabla\Psi = -\Delta B. \tag{28}$$

We see that $\partial_t h = 0$ implies $\Phi = Cst$, and $\partial_t q = 0$ implies $J(\Psi,q) = 0$, which means that isolines of potential vorticity are the mass transport streamlines. This is the case if for instance $q = F(\Psi)$. Reciprocally, if $\Phi = Cst$ and if q and Ψ have the same isolines, then $\partial_t h = 0$ and $\partial_t q = 0$, which also implies $\partial_t\Psi = 0$. Thus, through the decomposition (21), the velocity field is also stationary, with $\mathbf{u} = (1/h)\nabla^{\perp}\Psi$. We conclude that a necessary and sufficient condition for a shallow water flow $(h,\mathbf{u})$ to be stationary is

$$\Phi = Cst, \quad q \text{ and } \Psi \text{ have the same isolines.} \tag{29}$$

There is an additional relation verified by the stationary flow. This relation may be obtained by considering stationarity of the kinetic energy. The dynamics of the kinetic energy is obtained from Eq. (12):

$$\partial_t\left(\frac{\mathbf{u}^2}{2}\right) = -\frac{1}{h}J(\Psi,B) - \frac{1}{h}\nabla\Phi\cdot\nabla B. \tag{30}$$

Stationarity of the kinetic energy field and Eq. (29) gives $J(\Psi,B) = 0$: the Bernoulli function defined in Eq. (8) and the mass transport streamfunction Ψ have the same isolines. In addition, in any subdomain where $q = F(\Psi)$, the stationarity of the velocity field and the momentum equations (12) give the relation

$$q = \frac{\mathrm{d}B}{\mathrm{d}\Psi}. \tag{31}$$

2.3 Conserved quantities

Provided that the velocity and height fields remain differentiable, the shallow water dynamics in Eqs. (6), (13) and (14) conserves the total energy

$$\mathscr{E}[\mathbf{u},h] = \frac{1}{2}\int \mathrm{d}\mathbf{x}\ \left[h\mathbf{u}^2 + g\left(h + h_b - 1\right)^2\right], \tag{32}$$

which includes a kinetic energy contribution and a potential energy contribution. It is known that the shallow water dynamics sometimes leads to shocks that prevent energy conservation. We postpone a discussion of this important point to the last section of this paper.

As a consequence of particle relabeling symmetry there is an infinite number of other conserved quantities called Casimir functionals, see e.g. [35]. These functionals are written

$$\mathscr{C}_s = \int \mathrm{d}\mathbf{x}\, hs(q) \;, \tag{33}$$

where s is any distribution. This conservation can be easily checked from Eqs. (13) and (6). The conservation of all the Casimirs implies the conservation of all the potential vorticity moments

$$\forall k \in \mathbb{N},\; \mathscr{Z}_k = \int \mathrm{d}\mathbf{x}\, hq^k \;. \tag{34}$$

These Casimir functionals include the total mass conservation ($k = 0$), the conservation of the circulation[2] ($k = 1$) and the conservation of the enstrophy ($k = 2$).

The conservation of all the Casimirs is equivalent to the conservation of mass plus the conservation of the potential vorticity distribution $\mathscr{D}[q]$ defined through

$$\forall \sigma, \quad \mathscr{D}[q](\sigma)\,\mathrm{d}\sigma = \int \mathrm{d}\mathbf{x}\, h\mathscr{I}_{\{\sigma \le q \le \sigma + \mathrm{d}\sigma\}} \tag{35}$$

where $\mathscr{I}_{\{\sigma \le q \le \sigma + \mathrm{d}\sigma\}}$ is the characteristic function, i.e. it returns one if $\sigma \le q(\mathbf{x}) \le \sigma + \mathrm{d}\sigma$ and zero otherwise. It means that the global volume of each potential vorticity level σ_q is conserved through the dynamics.

We will restrict ourselves to initial states where the potential vorticity field has a distribution characterized by its moments only, then the knowledge of its global distribution given in Eq. (35) and of the total mass is equivalent to the knowledge of the moments given in Eq. (34).

Depending on the domain geometry, there could be additional invariants. For the sake of simplicity, we do not discuss the role of these additional invariant in this paper, but it would not be difficult to generalize our results by taking them into account.

In the quasi-geostrophic or two-dimensional Euler models, dynamical invariants have important consequences such as the large scale energy transfer, the self-organization at the domain scale, and the existence of an infinite number of stable states for the dynamics, see for instance Ref. [4] and references therein. We show in the present paper that these dynamical invariants play a similar role in the shallow water case, allowing for a large-scale circulation associated with the potential vorticity field, even if this process may be associated with a concomitant loss of energy toward small scales.

3 Equilibrium statistical mechanics of a discrete shallow water model

The aim of this section is to compute statistical equilibrium states of a discrete model of the shallow water system and to consider the thermodynamic limit for this model. In the first subsection, a formal Liouville theorem is given for the triplet of variables (h, hu, hv) and then the triplet (h, q, μ) after a change of variables. The derivation given with more details in Appendix A is made in the Eulerian representation. This allows to write formally the microcanonical measure of the shallow water model for these triplets of fields.

In the second subsection, a finite dimensional semi-Lagrangian discretization of the model is proposed to give a physical meaning to the formal measure. Other discretizations of the shallow water

[2] The actual circulation is usually defined as $\Gamma \equiv \int_{\partial D} \mathrm{d}\mathbf{l} \cdot \mathbf{u}$, where $\mathrm{d}\mathbf{l}$ is a vector tangent to the domain boundary. Stokes theorem yields to $Z_1 = \Gamma + f$.

system may have been considered but any other choices we have tried were not taking into account the fluid particle mass conservation and led to inconsistent results (for instance the equilibrium states were not stationary, not stable by coarse-graining, and unbalanced, see Appendix E).

Then the macroscopic states are defined through a coarse-graining procedure in a third subsection. The variational problem leading to the equilibrium states, i.e. the most probable macroscopic state, is introduced in the fourth subsection. This variational problem generalizes the Miller-Robert-Sommeria equilibrium theory to the shallow water system.

3.1 Liouville theorem

The first step before computing equilibrium states is to define what is a microscopic configuration of the system, which requires to identify the relevant phase space. The simplest set of variables to consider are those that satisfy a Liouville theorem. A Liouville theorem means that the flow in phase space is non-divergent, which implies the time invariance of the microcanonical ensemble. The existence of a Liouville theorem for the shallow water system was initially shown by Warn [53] who considered a decomposition of the flow fields on a basis given by the eigenmodes of the linearized dynamics. A Liouville theorem was shown by Weichman and Petrich[56] for a Lagrangian representation of the dynamics. Following a general method proposed by one of us and described by Thalabard [40], it is shown in Appendix A that the triplet (h, hu, hv) does satisfy a formal Liouville theorem. At a formal level, the microcanonical measure can then be written[3]

$$\mathrm{d}\mu_{h,hu,hv}\left(E,\{Z_k\}_{k\geq 0}\right)=\frac{1}{\Omega\left(E,\{Z_k\}_{k\geq 0}\right)}\mathscr{D}[h]\,\mathscr{D}[hu]\,\mathscr{D}[hv]\,\delta\left(\mathscr{E}-E\right)\prod_{k=0}^{+\infty}\delta\left(\mathscr{Z}_k-Z_k\right), \tag{36}$$

with the phase space volume

$$\Omega\left(E,\{Z_k\}_{k\geq 0}\right)=\int\mathscr{D}[h]\,\mathscr{D}[hu]\,\mathscr{D}[hv]\,\delta\left(\mathscr{E}-E\right)\prod_{k=0}^{+\infty}\delta\left(\mathscr{Z}_k-Z_k\right). \tag{37}$$

Here $\mathscr{E}$ is the energy of a microscopic configuration defined in Eq. (32), and the $\{\mathscr{Z}_k\}_{k\geq 0}$ are the potential vorticity moments of a microscopic configuration defined in Eq. (34). Those constraints are the dynamical invariants of the shallow water model. The notation $\int\mathscr{D}[h]\,\mathscr{D}[hu]\,\mathscr{D}[hv]$ means that the integral is formally performed over each possible triplet of fields $(h,\ hu,\ hv)$. The microcanonical measure allows to compute the expectation of an observable $A[h,hu,hv]$ in the microcanonical ensemble as

$$\langle\mathscr{A}\rangle_{d\mu_{h,hu,hv}}=\int\mathrm{d}\mu_{h,hu,hv}\ \mathscr{A}[h,hu,hv]\,. \tag{38}$$

Assuming ergodicity, the ensemble average $\langle\mathscr{A}\rangle_{d\mu_{h,hu,hv}}$ can finally be interpreted as the time average of the observable $\mathscr{A}$.

The triplet (h,hu,hv) is not a convenient one to work with, since the Casimir functionals $\{\mathscr{Z}_k\}_{k\geq 0}$ defined in Eq. (34) are not easily expressed in terms of these fields. Indeed, the expression of the Casimir functionals $\{\mathscr{Z}_k\}_{k\geq 0}$ involve not only the triplet (h,hu,hv), but also the triplet of the horizontal derivatives of these fields. We showed in subsection 2.1 that the triplet of fields (h,q,μ) fully describes the shallow water dynamics in a closed domain, and the functionals $\{\mathscr{Z}_k\}_{k\geq 0}$ are much more easily expressed in terms of the fields h,q. It is therefore more convenient to use these fields as independent variables. The price to pay is that the simple form of the energy defined in Eq. (32) in terms of the triplet (h,hu,vh) becomes more complicated when expressed in terms of the triplet (h,q,μ). However,

[3] The letter "μ" appearing in the measure denoted $\mathrm{d}\mu$ is not related to the divergent field denoted μ.

we will propose a simplified version of this energy functional, and argue in Appendix B that this is the relevant form of the energy to consider to compute the equilibrium state.

Unfortunately no direct proof of a Liouville theorem can be obtained for the triplet (h,q,μ). However, it is still possible to start from the microcanonical measure built with (h,hu,hv) in Eq. (36), and change variables at a formal level.It is shown in Appendix A.2 that the Jacobian of the transformation is[4]

$$\left|J\left[\frac{(h,hu,hv)}{(q,h,\mu)}\right]\right| = Ch^3. \tag{39}$$

where C is a constant. We conclude that the microcanonical measure can therefore be formally written

$$\mathrm{d}\mu_{h,q,\mu}\left(E,\{Z_k\}_{k\geq 0}\right) = \frac{1}{\Omega\left(E,\{Z_k\}_{k\geq 0}\right)} h^3 \mathscr{D}[h]\,\mathscr{D}[q]\,\mathscr{D}[\mu]\,\delta(\mathscr{E}-E)\prod_{k=0}^{+\infty}\delta(\mathscr{Z}_k - Z_k), \tag{40}$$

with

$$\Omega\left(E,\{Z_k\}_{k\geq 0}\right) = \int h^3 \mathscr{D}[h]\,\mathscr{D}[q]\,\mathscr{D}[\mu]\,\delta(\mathscr{E}-E)\prod_{k=0}^{+\infty}\delta(\mathscr{Z}_k - Z_k). \tag{41}$$

In that case, the expectation of an observable $A[q,h,\mu]$ in the microcanonical ensemble is given by

$$\langle \mathscr{A} \rangle_{d\mu_{q,h,\mu}} = \int \mathrm{d}\mu_{q,h,\mu}\mathscr{A}[q,h,\mu]. \tag{42}$$

A finite-dimensional projection of the fields will be given in the next subsection to give a meaning to the formal notations of this subsection.

3.2 A discrete model

In this subsection, we devise a discrete model of the shallow water system based on a semi-Lagrangian representation. This allows to give a finite dimensional representation of the formal measure given Eq. (40).

In order to keep track of the conservation properties of the continuous dynamics, we propose a convenient discretization of the shallow water system in terms of fluid particles. Since the fluid is considered incompressible, it is discretized into equal volume particles. The fluid particles do not have necessarily the same height, and the horizontal velocity field may be divergent. It is not possible to build a uniform grid with a single fluid particle per grid point since two particles with equal volume and different height can not occupy the same area. In order to bypass this difficulty, we define a uniform grid where several fluid particles can occupy a given site. To keep track that the actual fluid is continuous and incompressible, we add the condition that the area occupied by the particles inside a grid site fit the available area of the grid site.

In a first step the ensemble of microscopic configurations (the ensemble of the microstates of the discrete model) is defined. In a second step, the constraints of the discrete shallow water system are introduced, which allows to define the microcanonical measure of the discrete model in a third step.

[4] The h^3 term that appears after the change of variables in the functional integral must be understood as a "functional product" $\prod_{\mathbf{x}\in\mathscr{D}} h^3(\mathbf{x})$, see also the finite-dimensional representation of this measure given in the next subsection.

3.2.1 Definition of the ensemble of microscopic configurations

For the sake of simplicity, the domain $\mathscr{D}$ where the flow takes place is considered rectangular but generalizing the results to any shape would be straightforward. We recall that the horizontal and vertical length units have been chosen such that the domain area and the mean height are equal to one ($|\mathscr{D}| = L_x L_y = 1$, $H = 1$). The domain $\mathscr{D}$ is discretized into a uniform grid with $N = N_x \times N_y$ sites. The area of a grid site is $|\mathscr{D}|/N = 1/N$. Each site can contain many fluid particles. It is assumed that the fluid contains $N \times M$ fluid particles of equal mass and volume (the fluid is incompressible), where M is the average number of particles per site. The volume of a fluid particle is therefore

$$\delta V = \frac{|\mathscr{D}|H}{NM} = \frac{1}{NM}. \tag{43}$$

The grid sites are indexed by (i, j) with $1 \le i, j \le N_x, N_y$ and the fluid particles by n with $1 \le n \le NM$.

Each fluid particle is characterized by its position (I_n, J_n) on the grid, by its potential vorticity $q_n \in [q_{min}, q_{max}]$, its divergence $\mu_n \in [-\mu_{max}, \mu_{max}]$ and its height $h_n \in [h_{min}, h_{max}]$. The cutoffs on the potential vorticity can be physically related to actual minimum and maximum in the global distribution of potential vorticity level defined in Eq. (35), since this distribution is conserved by the dynamics. Such a justification does not exist for the other cutoffs. We will consider first the limit of an infinite number of fluid particles per grid site ($M \to +\infty$), then the limit of an infinite number of grid site ($N \to \infty$) and finally the limit of infinite height and divergence cut-off $\mu_{max} \to +\infty$, $h_{max} \to +\infty$, $h_{min} \to 0$. We will see that the result does not depend on those cut-off.

Let us introduce M_{ij} the number of particles per grid site (i, j), defined as

$$M_{ij} = \sum_{n=1}^{NM} \delta_{I_n,i}\delta_{J_n,j}. \tag{44}$$

The set of the particles that belong to the site (i, j) is denoted

$$\mathscr{M}_{ij} = \{1 \le n \le NM \,|\, (I_n, J_n) = (i, j)\}, \tag{45}$$

whose cardinal is M_{ij}. Mass conservation states that the total number of particles filling the grid is a constant, which gives the constraint

$$\sum_{ij} M_{ij} = NM, \tag{46}$$

where $\sum_{ij}$ means that we sum over the all the sites of the grid. A fluid particle labeled by n and carrying the height h_n occupies an area $\delta V/h_n$. The constraint that the area of each grid site is covered by fluid particles leads to the constraint

$$\forall i, j \qquad \frac{1}{M} \sum_{n \in \mathscr{M}_{ij}} \frac{1}{h_n} = 1, \tag{47}$$

where $\mathscr{M}_{ij}$ is the set defined in Eq. (45).

The ensemble of microstates of the discrete model is given by the set of all reachable values of grid positions, potential vorticities, divergences and heights of each fluid particle in accordance with the constraint of particle filling the area of each grid site:

$$X_{micro} \equiv \left\{ \chi_{micro} = \{(I_n, J_n), q_n, \mu_n, h_n\}_{1 \le n \le MN} \,\middle|\, \forall i, j \;1 \le i, j \le N_x, N_y \quad \frac{1}{M} \sum_{n \in \mathscr{M}_{ij}} \frac{1}{h_n} = 1 \right\}. \tag{48}$$

3.2.2 Coarse-graining

Here we consider a microstate $\chi_{micro} = \{(I_n, J_n), q_n, \mu_n, h_n\}_{1\leq n\leq MN}$ and an arbitrary function

$$g : n \to g_n = g(h_n, q_n, \mu_n). \tag{49}$$

We introduce two different coarse-graining procedures: an areal coarse-graining, and a volumetric coarse-graining.

The *areal coarse-graining* of the function g is defined at each grid point (i, j) as

$$\overline{g}_{ij} \equiv \frac{1}{M} \sum_{n\in\mathscr{M}_{ij}} \frac{1}{h_n} g_n \tag{50}$$

where $\mathscr{M}_{ij}$ is a set defined in Eq. (45). The terms $1/h_n$ appearing in Eq. (50) means that we consider local average of g_n weighted by the area occupied by each fluid particle. Note that we will only consider function g such that $\overline{g}_{ij}$ converges to a finite value in the limit of large $M_{ij} \sim M$. This means that the terms g_n should not be allowed to scale with M. This is the reason why we will consider first the large M limit, and then the limit of large cut-off μ_{max} and h_{max} for the fields μ and h.

The area filling constraint in Eq. (47) can then be written in terms of this areal coarse graining:

$$\overline{1}_{ij} = 1. \tag{51}$$

for any i, j. We also notice that the areal coarse-grained height field is simply the ratio of the number of particles in the site (i, j) over the averaged number of particles per site:

$$\overline{h}_{ij} = \frac{M_{ij}}{M}. \tag{52}$$

The *volumetric coarse-graining* of the function g is defined at each grid point (i, j) as

$$\langle g\rangle_{ij} \equiv \frac{1}{M_{ij}} \sum_{n\in\mathscr{M}_{ij}} g_n. \tag{53}$$

This field corresponds to the average of the function g carried by a fluid particle on site (i, j). The volumetric coarse-graining is related to the areal one through

$$\langle g\rangle_{ij} = \frac{\overline{hg}_{ij}}{\overline{h}_{ij}}. \tag{54}$$

3.2.3 Definition of a velocity field on the grid

Let us now define a large scale velocity field (or mean flow) on the uniform grid of the discrete model. We will introduce later a field accounting for small scale fluctuations of the velocity at each grid point. In the case of the actual shallow water model, for a given triplet of continuous fields h, q, μ, the velocity field is computed by using Eq. (20), which involves two spatial differential operators, namely $\nabla^\perp\Delta^{-1}$ and $\nabla\Delta^{-1/2}$. Discrete approximations of these spatial operators are well defined on the uniform grid of the discrete model. Discrete approximations of Eq. (20) can therefore be used to define a velocity field on the uniform grid. Let us consider a coarse-grained vorticity field $\widetilde{\omega} = \overline{h}\langle q\rangle - f$ and a coarse-grained divergent field $\widetilde{\mu} = \langle\mu\rangle$ defined on the same uniform grid. Discrete approximations of the operators appearing in Eq. (20) can be written respectively as $\left\{\nabla^\perp\Delta^{-1}[\widetilde{\omega}]\right\}_{ij} = \sum_{kl}\mathbf{G}^{\omega}_{ij,kl}.\widetilde{\omega}_{kl}$ and $\left\{\nabla^\perp\Delta^{-1}[\widetilde{\mu}]\right\}_{ij} = \sum_{kl}\mathbf{G}^{\mu}_{ij,kl}.\widetilde{\mu}_{kl}$, where the sum $\sum_{kl}$ is performed over each grid site (k, l). In the remainder of this paper, we do not need the explicit expression of the kernels $\{\mathbf{G}^{\omega}\}$ and $\{\mathbf{G}^{\mu}\}$, which depend only on the domain geometry. Using these notations, we define the large scale velocity field as

$$\mathbf{u}_{mf,ij} = \left\{\nabla^{\perp}\Delta^{-1}\left[\overline{h}\langle q\rangle - f\right]\right\}_{ij} + \left\{\nabla\Delta^{-1/2}\left[\langle\mu\rangle\right]\right\}_{ij}, \tag{55}$$

where the index "mf" stands for "mean flow".

At this point one may wonder why the relevant coarse-grained fields used to define the large scale flow $\mathbf{u}_{mf}$ should be $\overline{h}, \langle\mu\rangle, \langle q\rangle$ (which, using Eq. (54) , is equivalent to either the triplet $\overline{h}, \overline{h\mu}, \overline{hq}$ or to the triplet $\left\langle h^{-1}\right\rangle, \langle\mu\rangle, \langle q\rangle$). Our motivation for such a choice is twofold. First, we will see *a posteriori* that this allows to recover previous results derived in several limit cases (weak flow limit, quasi-geostrophic limit), and to obtain consistent results in the general case. Second, in physical space, the number of fluid particles at each grid site is given by the areal-coarse-grained height field $\overline{h}$, according to the previous section. The relevant macroscopic potential vorticity field or divergent field is then given by the volumetric coarse-graining $\langle q\rangle = \overline{hq}/\overline{h}$ and $\langle\mu\rangle = \overline{h\mu}/\overline{h}$. In that respect, the phase space variable h does not play the same role as q and μ when considering macroscopic quantities in physical space, and this is why we do not consider the triplet $\langle h\rangle, \langle\mu\rangle, \langle q\rangle$ to describe the system at a macroscopic level.

3.2.4 Definition of the microcanonical ensemble for the discrete model.

Here we introduce a set of constraints associated with the discrete model in order to define the microcanonical ensemble. These constraints are a discrete version of the Casimirs functional and the energy of the continuous shallow water model, defined in Eqs. (34) and (32), respectively. Additional assumptions on the form of the energy will be required, and we will discuss the relevance of such assumptions.

Note that we introduce here constraints for the discrete shallow water model, but we do not define what would be the dynamics of the discrete model. Indeed, it is not necessary to know the dynamics in order to compute the equilibrium state of the system. Only the knowledge of constraints provided by dynamical invariants is required.

Let us consider a given microstates $\chi_{micro} = \{(I_n, J_n), q_n, \mu_n, h_n\}_{1\le n\le MN}$, which belongs to the ensemble X_{micro} of possible configurations of the discrete model defined in Eq. (48). By construction of the ensemble X_{micro}, each element χ_{micro} satisfies the areal filling constraint given in Eq. (47).

The potential vorticity moments of the discrete model are defined as

$$\forall k \ge 0, \quad \mathscr{Z}_k \equiv \frac{1}{NM}\sum_{n=1}^{NM} q_n^k = \frac{1}{N}\sum_{ij}\overline{hq^k}_{ij}. \tag{56}$$

The notation $\sum_{ij}$ means that the sum is performed over each grid point, with $1 \le i \le N_x$, $1 \le j \le N_y$. It is shown in Appendix B that those discrete potential vorticity moments tend to the potential vorticity moments of the continuous dynamics defined in Eq. (34) in the limit of large number of fluid particles NM.

We also define the total energy of the discrete model as

$$\mathscr{E} \equiv \frac{1}{NM}\sum_{n=1}^{NM}\left(\frac{1}{2}\mathbf{u}_{mf,I_nJ_n}^2 + \frac{1}{2}\left(\mu_n - \langle\mu\rangle_{I_nJ_n}\right)^2 + g\left(\frac{h_n}{2} + h_{bI_nJ_n}\right)\right) - \mathscr{E}_{cst}. \tag{57}$$

where $h_{b,ij}$ is the discrete topography[5], where $\mathbf{u}_{mf,ij}$ is the mean flow defined in Eq. (55), and where $\mathscr{E}_{cst}\left[h_{bij}\right]$ is an unimportant functional of h_b chosen such that $\mathscr{E} = 0$ at rest (i.e. when $\mu_n = 0$, $\mathbf{u}_{mfI_nJ_n} =$

[5] Here, the bottom topography is assumed sufficiently smooth to be considered constant over a grid site. A fluctuating topography would require further discussion.

0, $h_n = 1 - h_{bI_nJ_n}$ for any n and $M_{ij} = M$ for any (i,j)):

$$\mathscr{E}_{Cst} = \frac{1}{N}\sum_{ij}\frac{g}{2}\left(1 - h_{bij}^2\right). \tag{58}$$

A simple interpretation for this form of the total energy is that each fluid particle carries a kinetic energy associated with the mean flow $\mathbf{u}_{mf}$, as well as a kinetic energy associated with local fluctuations of the divergence field and finally a potential energy (the height of the center of mass of the fluid particle is $h_{bI_nJ_n} + h_n/2$).

It is argued in Appendix B that with only a few reasonable assumptions on the properties of the equilibrium state, the energy of the discrete model defined in Eq. (57) would also be the energy of the equilibrium state of the actual shallow water model defined in Eq. (32) in the limit of large number of fluid particles NM. Note also that according to Eq. (57), the vortical part of the velocity field does not contribute to local small scale kinetic energy, which is analogous to previous statistical mechanics results for non-divergent flow models such as two-dimensional Euler equations or quasi-geostrophic equations [23,33]. Qualitatively, this is due to the fact that inverting the Laplacian operator smooth out local fluctuations of the relative vorticity $\omega = hq - 1$ so that the streamfunction associated with the microscopic vorticity field is the same as the streamfunction associated with the coarse-grained relative vorticity field, see Appendix B for more details.

The expression of the energy in Eq. (57) involves a sum over the NM fluid particles. This sum can be recast into a sum over the N points of the grid, by using the definition of the areal coarse-graining in Eq. (50) and the definition of volumetric coarse-graining in Eq. (53):

$$\mathscr{E} = \frac{1}{N}\sum_{ij}\left(\frac{1}{2}\overline{h}_{ij}\mathbf{u}_{mf,ij}^2 + \frac{1}{2}\overline{h}_{ij}\left(\left\langle\mu^2\right\rangle_{ij} - \langle\mu\rangle_{ij}^2\right) + \frac{g}{2}\left(\left(\overline{h}_{ij} + h_{b,ij} - 1\right)^2 + \left(\overline{h^2}_{ij} - \overline{h}_{ij}^2\right)\right)\right). \tag{59}$$

This Eulerian representation of the energy allows to identify three different contributions.

One first contribution to the total energy is given by the sum over each grid point (i,j) of the kinetic energy of the mean flow $\mathbf{u}_{mf}$ (which is carried by $M_{ij} = M\overline{h}_{ij}$ fluid particles), and of the potential energy of the areal coarse-grained height field $\overline{h}_{ij}$:

$$\mathscr{E}_{mf} \equiv \frac{1}{2N}\sum_{ij}\left[\overline{h}_{ij}\mathbf{u}_{mf,ij}^2 + g\left(\overline{h}_{ij} + h_{b,ij} - 1\right)^2\right]. \tag{60}$$

This contribution will be referred to as the total mean flow energy, or the energy of the large scale flow.

A second contribution to the total energy is given by the sum over each grid points of the variance of the divergence levels μ carried by fluid particles at site (i,j) , times the number of fluid particles ($M_{ij} = M\overline{h}_{ij}$):

$$\mathscr{E}_{\delta\mu} \equiv \frac{1}{2N}\sum_{ij}\overline{h}_{ij}\left(\left\langle\mu^2\right\rangle_{ij} - \langle\mu\rangle_{ij}^2\right). \tag{61}$$

This term can be interpreted as a subgrid-scale (or small scale) kinetic energy term due entirely to the divergent part of the velocity field (see Appendix B).

The last contribution to the total energy is the sum over each grid point of the potential energy associated with local fluctuations of the height field:

$$\mathscr{E}_{\delta h} \equiv \frac{g}{2N}\sum_{ij}\left(\overline{h^2}_{ij} - \overline{h}_{ij}^2\right). \tag{62}$$

This term can be interpreted as a subgrid-scale (or small scale) potential energy term.

One can finally check that total energy defined in Eq. (59) is the sum of the three contributions given in Eqs. (60), (61) and (62) :

$$\mathscr{E} = \mathscr{E}_{mf} + \mathscr{E}_{\delta\mu} + \mathscr{E}_{\delta h}. \tag{63}$$

It is well known that for the 2D Euler model or the quasi-geostrophic model, there is no contribution to the energy from the sub-grid fluctuations of the potential vorticity in the limit of vanishing grid size (see [22,33]). For the shallow water model, this is also the case. However the sub-grid fluctuations of height and divergence do contribute to the energy (see Eqs. (61), (62) and (63)). A qualitative reason for this contribution of local height and divergence fluctuations to the total energy is that those two fields may be decomposed on the basis of inertia-gravity waves, which are known to develop a small scale energy transfer [53], and consequently to a loss of energy at subgrid-scales in the discretized model.

Now that we have defined the configurations space in Eq. (48), the potential vorticity moments $\{\mathscr{Z}_k[\chi_{micro}]\}_{k\geq 0}$ of the discrete model in Eq. (56) and the energy $\mathscr{E}[\chi_{micro}]$ of the discrete model in Eq. (59), we introduce the microcanonical ensemble as the restriction of the configurations space X_{micro} to configurations with fixed values of energy E and Casimirs $\{Z_k\}_{k\geq 0}$:

$$\left\{\chi_{micro} = \{(I_n,J_n), q_n, \mu_n, h_n\}_{1\leq n\leq MN} \in X_{micro} \,|\, \mathscr{E}[\chi_{micro}] = E, \forall k \in \mathbb{N} \quad \mathscr{Z}_k[\chi_{micro}] = Z_k\right\}. \tag{64}$$

In our discrete model, we assume that the microcanonical measure over the ensemble of configurations X_{micro} is :

$$\mathrm{d}\mu^{N,M}_{E,\{Z_k\}_{k\geq 0}}(\chi_{\mathrm{micro}}) = \frac{\delta(\mathscr{E}[\chi_{\mathrm{micro}}]-E)}{\Omega(E,\{Z_k\}_{k\geq 0})}\prod_{k=0}^{\infty}\delta(\mathscr{Z}_k[\chi\mathrm{micro}]-Z_k)\prod_{ij}\delta(\overline{1}_{ij}[\chi\mathrm{micro}]-1)$$
$$\times\prod_{n=1}^{NM} h_n^3\mathrm{d}h_n\mathrm{d}q_n\mathrm{d}\mu_n\mathrm{d}I_n\mathrm{d}J_n, \tag{65}$$

where $\Omega(E,\{Z_k\})$ is the phase space volume defined as

$$\Omega(E,\{Z_k\}_{k\geq 0}) = \sum_{I_1=1}^{N_x}\cdots\sum_{I_n=1}^{N_x}\sum_{J_1=1}^{N_y}\cdots\sum_{J_n=1}^{N_y}\int\left[\prod_{n=1}^{NM}h_n^3 dh_n dq_n d\mu_n\right]\delta(\mathscr{E}[\chi_{\mathrm{micro}}]-E)\prod_{k=0}^{\infty}\delta(\mathscr{Z}_k[\chi_{\mathrm{micro}}]-Z_k)$$
$$\times\prod_{ij}\delta(\overline{1}_{ij}[\chi_{\mathrm{micro}}]-1). \tag{66}$$

The terms $\mathrm{d}I_n$ and $\mathrm{d}J_n$ are discrete measures with support on the grid coordinates: $\mathrm{d}I_n(f) = \sum_{i=1}^{N_x} f_i$, $\mathrm{d}J_n(f) = \sum_{j=1}^{N_y} f_j$. The product $\prod_{ij}$ is performed over the grid sites (i,j) with $1\leq i\leq N_x$ and $1\leq j\leq N_y$. The constraint $\delta\left(\overline{1}_{ij}[\chi_{micro}]-1\right)$ corresponds to the area filling constraint defined in Eq. (51), which must be satisfied by each microstate $\chi_{micro}\in X_{micro}$.

Note that Eq. (65) is a discrete version of the formal microcanonical measure given in Eq. (40) for the continuous case. The expectation of an observable $\mathscr{A}[\chi_{micro}]$ in the microcanonical ensemble is

$$\langle\mathscr{A}\rangle^{N,M}_{E,\{Z_k\}_{k\geq 0}} = \mathrm{d}\mu^{N,M}_{E,\{Z_k\}_{k\geq 0}}(\mathscr{A}) = \int \mathrm{d}\mu^{N,M}_{E,\{Z_k\}_{k\geq 0}}(\chi_{\mathrm{micro}})\mathscr{A}[\chi_{\mathrm{micro}}]. \tag{67}$$

The problem is now to compute the macrostate entropy, which accounts for the logarithm of the number of microstates corresponding to the same macrostate. Although we do not need to consider this problem in the present work, it would be very interesting to have continuous or discrete approximation of the shallow water equation that have the invariant measure (65), in an analogous way as what was achieved for the 2D Euler equations [9,13].

3.3 Macrostates and their entropy

The aim of this subsection is to compute the equilibrium state of the discrete model introduced in the previous subsection. The first step is to define the macrostates of the system, the next step is to compute the most probable macrostates, using large deviation theory, which yields also a concentration property asymptotically for large N and M (almost all the microstates correspond to the most probable macrostate).

3.3.1 Definition of the empirical density field

Let us consider a microstate $\chi_{micro} = \{(I_n, J_n), q_n, \mu_n, h_n\}_{1 \le n \le MN}$ picked in the ensemble of configurations X_{micro} defined in Eq. (48). The empirical density field of this microstate is defined for each grid point (i, j) as

$$p_{ij}\left(\sigma_h, \sigma_q, \sigma_\mu\right)[\chi_{\text{micro}}] \equiv \overline{\delta\left(h - \sigma_h\right)\delta\left(q - \sigma_q\right)\delta\left(\mu - \sigma_\mu\right)}_{ij}, \tag{68}$$

where the overline operator is the areal coarse-graining defined in (50). This field contains all the statistical information of the system at the grid level. The constraint that each grid site (i, j) is covered by particles is given by Eq. (51), which ensures the normalization:

$$\int \mathrm{d}\sigma_h \mathrm{d}\sigma_q \mathrm{d}\sigma_\mu \, p_{ij} = 1. \tag{69}$$

Let us consider a function $g(h_n, q_n, \mu_n)$ depending on the height, potential vorticity and divergence carried by a fluid particle. Let us then consider the discrete microscopic field $n \to g_n = g(h_n, q_n, \mu_n)$ with $1 \le n \le NM$. Following Eqs. (50) and (68), the corresponding coarse-grained field $\overline{g}_{ij}$ is expressed solely in terms of the empirical density field p_{ij}:

$$\overline{g}_{ij} = \int \mathrm{d}\sigma_h \mathrm{d}\sigma_q \mathrm{d}\sigma_\mu \, g(\sigma_h, \sigma_q, \sigma_\mu) p_{ij}\left(\sigma_h, \sigma_q, \sigma_\mu\right). \tag{70}$$

If we consider for instance the function $g(h, q, \mu) = q$, then $g_n = q_n$ the microscopic potential vorticity field, and coarse-grained potential vorticity field $\overline{q}_{ij}$ is obtained by a direct application of Eq. (70): $\overline{q}_{ij} = \int \mathrm{d}\sigma_h \mathrm{d}\sigma_q \mathrm{d}\sigma_\mu \, \sigma_q p_{ij}$.

Importantly, the constraints $\{\mathscr{Z}_k[\chi_{micro}] = Z_k\}_{k \ge 0}$ and $\mathscr{E}[\chi_{micro}] = E$ defined in Eq. (56) and (59) depend only on the empirical density field p_{ij} (since they depend only on local areal coarse-grained moments of the different fields). The empirical density is therefore a relevant variable to fully characterize the system at a macroscopic level.

3.3.2 Definition of the macrostates

The macrostates are defined as the set of microscopic configurations leading to a given value $p_{ij} = \rho_{ij}$ of the empirical density field:

$$\rho \equiv \left\{\chi_{micro} \in X_{micro} \,\middle|\, \forall i, j \quad p_{ij}[\chi_{micro}] = \rho_{ij}\right\}. \tag{71}$$

For the sake of simplicity, we make a small abuse of notation by denoting ρ both the macrostate defined in Eq. (71) and the field $\rho = \left\{\rho_{ij}\right\}$. The values of the constraints are the same for all microstates within a given macrostate since they depend only on the local coarse-grained moments of the different fields, which remain unchanged for a prescribed empirical density field. The energy and the Casimirs, defined in Eqs. (57) and (56) respectively, have the same values for all the microstates within a single macrostate and will therefore be denoted by $\mathscr{E}[\rho]$ and $\{\mathscr{Z}_k[\rho]\}_{k \ge 0}$.

3.3.3 Macroscopic observables and empirical density

Let us now consider an observable $\mathscr{A}[\chi_{micro}]$ on the configuration space X_{micro} defined in Eq. (48) such that their dependance on the microscopic configuration χ_{micro} occurs only through the empirical density field:

$$\mathscr{A}[\chi_{micro}] = \mathscr{A}\left[\left\{p_{ij}[\chi_{micro}]\right\}\right]. \tag{72}$$

This is actually the case for any observable written as a sum over the fluid particles, i.e. for any observable appearing in Eq. (67). It is therefore possible to change variables from χ_{micro} to the empirical density field values $\left\{\rho_{ij}\right\}_{1\le i,j\le N_x,N_y}$ in Eq. (67):

$$\langle\mathscr{A}\rangle^{N,M}_{E,\{Z_k\}_{k\ge0}} = \int\left[\prod_{ij}\mathscr{D}[\rho_{ij}]\right]\mathscr{A}\left[\left\{\rho_{ij}\right\}\right]\frac{\Omega(\rho)}{\Omega\left(E,\{Z_k\}_{k\ge0}\right)}\delta\left(\mathscr{E}[\rho]-E\right)\prod_{k=0}^{\infty}\delta\left(\mathscr{Z}_k[\rho]-Z_k\right) \times\prod_{ij}\delta\left(\overline{1}_{ij}[\rho_{ij}]-1\right), \tag{73}$$

where

$$\Omega(\rho) = \sum_{I_1=1}^{N_x}\cdots\sum_{I_n=1}^{N_x}\sum_{J_1=1}^{N_y}\cdots\sum_{J_n=1}^{N_y}\int\left[\prod_{n=1}^{NM}h_n^3 dh_n dq_n d\mu_n\right]\prod_{ij}\left[\hat{\delta}(p_{ij}[\chi_{\text{micro}}]-\rho_{ij})\right] \tag{74}$$

is the volume of a macrostate ρ defined in Eq. (71) in the configuration space X_{micro} defined in Eq. (48), and where $\Omega\left(E,\{Z_k\}_{k\ge0}\right)$ is the total volume in phase space defined in Eq. (66). The constraint $\delta\left(\overline{1}_{ij}[\rho_{ij}]-1\right)$ is a normalization constraint for the $\left\{\rho_{ij}\right\}$, since $\overline{1}_{ij}[\rho_{ij}] = \int \mathrm{d}\sigma_h\mathrm{d}\sigma_q\mathrm{d}\sigma_\mu\,\rho_{ij}$. This normalization constraint also corresponds to the constraint that the fluid particles within a site must fit the available area, see Eq. (47). The term $\int\mathscr{D}[\rho_{ij}]$ means that the integral is performed over all the possible functions ρ_{ij}. The term $\hat{\delta}(p_{ij}-\rho_{ij})$ is a Dirac delta distribution on the functional space of the empirical density field values $\rho_{ij}(\sigma_h,\sigma_q,\sigma_\mu)$, with $p_{ij}\left(\sigma_h,\sigma_q,\sigma_\mu\right)[\chi_{micro}]$ the empirical density field defined in Eq. (68).

3.3.4 Asymptotic behavior of the macrostates volume and derivation of its entropy

Let us compute the asymptotic form of $\Omega(\rho)$ defined in Eq. (74), by considering the limit $M\to\infty$, where M is the average number of particles per grid site (i,j). For a given set of macrostates ρ defined in Eq. (71), the number of fluid particles per grid site is

$$M_{ij} = M\int \mathrm{d}\sigma_h\mathrm{d}\sigma_q\mathrm{d}\sigma_\mu\,\sigma_h\rho_{ij}. \tag{75}$$

This is the only constraint on the particle positions in the grid for a microstate that belongs to the macrostates ρ. All the realizations of the particle positions that satisfies (75) count equally in the ensemble of macrostates ρ. Through combinatorial calculation, the number of realizations of $\{(I_n,J_n)\}$ that satisfy (75) is $(NM)!/\prod_{ij}M_{ij}!$. Using Eq. (52) to express M_{ij} in terms of $\overline{h}_{ij}$, we get

$$\Omega(\rho) = \frac{(NM)!}{\prod_{ij}\left(M\overline{h}_{ij}\right)!}\prod_{ij}\Omega_{ij}, \tag{76}$$

where

$$\Omega_{ij} \equiv \int\left[\prod_{m=1}^{M\overline{h}_{ij}}h_m^3\mathrm{d}h_m\mathrm{d}q_m\mathrm{d}\mu_m\right]\hat{\delta}\left(\frac{1}{M}\sum_{m=1}^{M\overline{h}_{ij}}\frac{1}{h_m}\delta\left(h_m-\sigma_h\right)\delta\left(q_m-\sigma_q\right)\delta\left(\mu_m-\sigma_\mu\right)-\rho_{ij}\right) \tag{77}$$

is the number of possible configurations for a given set of $M_{ij} = M\overline{h}_{ij}$ fluid particles at site (i,j).

The asymptotic behavior of the pre-factor in Eq. (76) is computed through the Stirling formula:

$$\log\left(\frac{(NM)!}{\prod_{ij}\left(M\overline{h}_{ij}\right)!}\right) \underset{M\to\infty}{\sim} MN\left(-\frac{1}{N}\sum_{ij}\overline{h}_{ij}\log\left(\overline{h}_{ij}\right)+\log(N)\right). \tag{78}$$

The asymptotic behavior with M of Ω_{ij} defined in Eq. (77) can be computed by using Sanov's theorem.

Before applying this theorem to our problem, let us consider the simpler case of K independent and identically distributed variables $\{\chi_k\}_{1\le k\le K}$ with common probability density function $F(\chi)$. Those variable take values in a bounded interval of $\mathbb{R}$. [6] Sanov's theorem describes the large deviation of the empirical density distribution

$$f_K \equiv \frac{1}{K}\sum_{k=1}^{K}\delta\left(\chi_k-\chi\right), \tag{79}$$

which can be considered as an actual probability distribution for the variable χ. The probability distribution functional of this empirical density function is

$$\mathscr{P}[f] \equiv \int\prod_{k=1}^{K}F\left(\chi_k\right)\mathrm{d}\chi_k\,\hat{\delta}\left(f-f_K\right). \tag{80}$$

Sanov's theorem is a statement about the asymptotic behavior of the logarithm of $\mathscr{P}[f]$. For a given function $f(\chi)$, Sanov's theorem states

$$\log\left(\mathscr{P}[f]\right) \underset{K\to+\infty}{\sim} -K\int\mathrm{d}\chi\, f(\chi)\log\left(\frac{f(\chi)}{F(\chi)}\right) \tag{81}$$

if $\int\mathrm{d}\chi\, f(\chi)=1$ and $\log\left(\mathscr{P}[f]\right)\sim-\infty$ otherwise. An heuristic discussion of Sanov's theorem is given in Ref [43].

Combining Eqs. (79), (80) and (81) and generalizing this result to K independent and identically distributed L-tuple of variables $\left\{\{\chi_{l,k}\}_{1\le l\le L}\right\}_{1\le k\le K}$ with common probability density function $F\left(\{\chi_l\}_{1\le l\le L}\right)$, Sanov's theorem is written in compact form as

$$\log\left(\int\prod_{k=1}^{K}F\left(\{\chi_{l,k}\}_{1\le l\le L}\right)\prod_{l=1}^{L}\mathrm{d}\chi_{l,k}\,\hat{\delta}\left(f-\frac{1}{K}\sum_{k=1}^{K}\prod_{l=1}^{L}\delta\left(\chi_{l,k}-\chi_l\right)\right)\right) \underset{K\to\infty}{\sim} -K\int\prod_{l=1}^{L}\mathrm{d}\chi_l\, f\log\left(\frac{f}{F}\right) \tag{82}$$

if $\int\prod_{l=1}^{L}\mathrm{d}\chi_l\, f\left(\{\chi_l\}_{1\le l\le L}\right)=1$.

Let us come back to the asymptotic behavior of Ω_{ij} defined in Eq. (77), in the large M limit. Before applying Sanov's theorem, it is needed to recast this equation into a form similar to the argument of the logarithm in the lhs of Eq. (82). Because of the $1/h_m$ term appearing in the delta Dirac function of Eq. (77), one needs to perform first a change of variable from ρ_{ij} to

$$\pi_{ij} \equiv \frac{\sigma_h}{\overline{h}_{ij}}\rho_{ij} \tag{83}$$

The formal Jacobian $\mathscr{J}$ arising from this change of variable in the functional delta Dirac function depend only on $\overline{h}_{ij}$ which does not depend on M and will therefore not matter for the asymptotic behavior of $\log\left(\Omega_{ij}\right)$. In addition, the factor h_m^3 appearing in front of the Lebesgue measure in Eq. (77)

[6] The fact that the variable χ_k have to be bounded is the reason why we set the cutoffs on the values of h_n, q_n and μ_n.

must be divided by a normalization factor $Z=\int h_m^3 \mathrm{d}h_m \mathrm{d}q_m \mathrm{d}\mu_m$ so that $P(h_m,q_m,\mu_m)=h_m^3/Z$ can be interpreted as a probability distribution function.[7] Then Ω_{ij} writes :

$$\Omega_{ij}=\mathscr{J} Z^{M\overline{h}_{ij}} \int \prod_{m=1}^{M\overline{h}_{ij}} \left(\frac{h_m^3}{Z}\right) \mathrm{d}h_m \mathrm{d}q_m \mathrm{d}\mu_m \, \hat{\delta}\left(\pi_{ij}-\frac{1}{M\overline{h}_{ij}} \sum_{m=1}^{M\overline{h}_{ij}} \delta\left(h_m-\sigma_h\right) \delta\left(q_m-\sigma_q\right) \delta\left(\mu_m-\sigma_\mu\right)\right). \tag{84}$$

A direct application of Sanov's theorem to Eq. (84) then yields

$$\log\left(\Omega_{ij}\right) \underset{M\to\infty}{\sim} M\overline{h}_{ij}\left[-\int \mathrm{d}\sigma_h \mathrm{d}\sigma_q \mathrm{d}\sigma_\mu \, \pi_{ij} \log\left(\frac{\pi_{ij}}{\sigma_h^3}\right)\right]. \tag{85}$$

The normalization constraint on the distributions π_{ij} is already fulfilled by the definition of the coarse-grained height fields $\overline{h}_{ij}=\overline{h}_{ij}\int \mathrm{d}\sigma_h \mathrm{d}\sigma_q \mathrm{d}\sigma_\mu \, \pi_{ij}$ (see Eq. (70)). The inverse change of variable $\rho_{ij}=\overline{h}_{ij}\pi_{ij}/\sigma_h$ in Eq. (85) yields

$$\log\left(\Omega_{ij}\right) \underset{M\to\infty}{\sim} M\left[-\int \mathrm{d}\sigma_h \mathrm{d}\sigma_q \mathrm{d}\sigma_\mu \, \sigma_h \rho_{ij} \log\left(\frac{\rho_{ij}}{\sigma_h^2}\right)+\overline{h}_{ij}\log\left(\overline{h}_{ij}\right)\right]. \tag{86}$$

Combining Eq. (76) with Eqs. (78) and (86), we obtain the asymptotic behavior of $\Omega\left(\rho\right)$ with M:

$$\log\left(\Omega\left(\rho\right)\right) \underset{M\to\infty}{\sim} MN\mathscr{S}\left[\rho\right], \tag{87}$$

where

$$\mathscr{S}\left[\rho\right]=-\frac{1}{N}\sum_{ij}\int \mathrm{d}\sigma_h \mathrm{d}\sigma_q \mathrm{d}\sigma_\mu \, \sigma_h \rho_{ij}\log\left(\frac{\rho_{ij}}{\sigma_h^2}\right) \tag{88}$$

is the macrostate entropy[8].

3.4 Continuous limit

3.4.1 Expressions of the macrostate entropy, energy and potential vorticity moments

Considering now the limit of an infinite number of grid site ($N\to\infty$), the site coordinates (i,j) tend toward the continuous space coordinates $\mathbf{x}$ and the macrostate entropy $\mathscr{S}$ derived in Eq. (88) becomes

$$\mathscr{S}\left[\rho\right]=-\int \mathrm{d}\mathbf{x}\mathrm{d}\sigma_h \mathrm{d}\sigma_q \mathrm{d}\sigma_\mu \, \sigma_h \rho\left(\mathbf{x},\sigma_h,\sigma_q,\sigma_\mu\right)\log\left(\frac{\rho\left(\mathbf{x},\sigma_h,\sigma_q,\sigma_\mu\right)}{\sigma_h^2}\right). \tag{89}$$

The empirical density has become a probability density function (pdf). For any function $g: n\to g_n=g\left(h_n,q_n,\mu_n\right)$, its continuous coarse-grained field is now computed through

$$\overline{g}\left(\mathbf{x}\right)=\int \mathrm{d}\sigma_h \mathrm{d}\sigma_q \mathrm{d}\sigma_\mu \, \rho\left(\mathbf{x},\sigma_h,\sigma_q,\sigma_\mu\right) g\left(\sigma_h,\sigma_q,\sigma_\mu\right), \tag{90}$$

The discrete mean flow defined in Eq. (55) becomes (by construction):

$$\mathbf{u}_{mf}\left[\overline{h},\overline{hq},\overline{h\mu}\right]=\nabla^\perp\Delta^{-1}\left(\overline{hq}-f\right)+\nabla\Delta^{-1/2}\left(\frac{\overline{h\mu}}{\overline{h}}\right), \tag{91}$$

[7] Here, Z depends on the cutoffs introduced in subsection 3.2. But we will see that it will vanish from the expression of the entropy in the end.

[8] Here we dropped the term $\log N$ coming from Eq. (78) as it is constant that can be discarded by redefining $\Omega_{E,\{Z_k\}_{k\geq 0}}$.

where all the coarse-grained fields have been expressed in terms of an areal coarse-graining, using Eq. (54). Similarly, the potential vorticity moments defined in Eq. (56) for the discrete model become

$$\forall k \in \mathbb{N} \quad \mathscr{Z}_k[\rho] = \int \mathrm{d}\mathbf{x}\, \overline{hq^k}. \tag{92}$$

The total energy defined in Eq. (59) for the discrete model becomes

$$\mathscr{E}[\rho] = \frac{1}{2}\int \mathrm{d}\mathbf{x}\, \left[\overline{h}\mathbf{u}_{mf}^2 + \left(\overline{h\mu^2} - \frac{\overline{h\mu}^2}{\overline{h}}\right) + g\overline{(h+h_b-1)^2}\right]. \tag{93}$$

We note that the total energy is a functional of $\overline{h}, \overline{h^2}, \overline{hq}, \overline{h\mu}, \overline{h\mu^2}$. Just as in the discrete case, this energy can be separated into a mean flow contribution (a large scale contribution including kinetic and potential energy), as well as a contribution from small scale kinetic energy due to local fluctuations of the divergent velocity field and a contribution from small scale potential energy due to local height fluctuations. The large scale (or mean flow) energy defined in Eq. (60) for the discrete model becomes

$$\mathscr{E}_{mf}\left[\overline{h}, \overline{hq}, \overline{h\mu}\right] = \frac{1}{2}\int \mathrm{d}\mathbf{x}\, \left[\overline{h}\mathbf{u}_{mf}^2 + g\left(\overline{h}+h_b-1\right)^2\right]. \tag{94}$$

The small scale (or subgrid-scale) kinetic energy due to local fluctuations of the divergent part of the velocity field defined in Eq. (61) becomes

$$\mathscr{E}_{\delta\mu}\left[\overline{h}, \overline{\mu}, \overline{\mu^2}\right] = \frac{1}{2}\int \mathrm{d}\mathbf{x}\, \left(\overline{h\mu^2} - \frac{\overline{h\mu}^2}{\overline{h}}\right), \tag{95}$$

and the small scale (or sub-grid scale) potential energy due to local height fluctuations defined in Eq. (62) becomes

$$\mathscr{E}_{\delta h}\left[\overline{h}, \overline{h^2}\right] = \frac{g}{2}\int \mathrm{d}\mathbf{x}\, \left(\overline{h^2} - \overline{h}^2\right). \tag{96}$$

One can check that the total energy in Eq. (93) is the sum of the three contributions given in Eq (94), (95) and (96):

$$\mathscr{E} = \mathscr{E}_{mf} + \mathscr{E}_{\delta\mu} + \mathscr{E}_{\delta h}. \tag{97}$$

3.4.2 Microcanonical variational problem for the probability density field

Let us come back to the average of a macroscopic observable $\mathscr{A}$ defined in Eq. (73). Using the asymptotic estimate for $\Omega[\rho]$ given in Eq. (87), the average of an observable $\mathscr{A}$ defined in Eq. (73) becomes[9]

$$\langle \mathscr{A} \rangle_{d\mu_{q,h,\mu}^{M,N}} = \int \mathscr{D}[\rho]\; \mathscr{A}[\rho] \frac{\mathrm{e}^{NM\mathscr{S}[\rho]}}{\Omega(E,\{Z_k\})} \delta(\mathscr{E}[\rho]-E) \prod_{k=0}^{\infty} \delta(\mathscr{Z}_k[\rho]-Z_k) \prod_{\mathbf{x}} \delta\left(\mathscr{I}[\rho]-1\right). \tag{98}$$

It comes to a Laplace-type integral where $NM \to \infty$. Thus, the value of $\langle \mathscr{A} \rangle_{d\mu_{q,h,\mu}^{M,N}}$ will be completely dominated by the contribution of the pdf ρ that maximizes the macrostate entropy defined in Eq. (89) while satisfying the normalization constraint

$$\mathscr{I}[\rho] = \int \mathrm{d}\sigma_h \mathrm{d}\sigma_q \mathrm{d}\sigma_\mu\, \rho\left(\mathbf{x}, \sigma_h, \sigma_q, \sigma_\mu\right) = 1, \tag{99}$$

[9] Strictly speaking, the equal sign should be noted $\asymp$ which means that the logarithm of the terms on both sides are equivalent, see e.g. Ref. [43].

and the microcanonical constraints $\mathscr{E}[\rho] = E$, $\{\mathscr{Z}_k[\rho] = Z_k\}_{k\geq 0}$, where the energy is defined in Eq. (93) and the potential vorticity moments are defined in Eq. (92). This variational problem can be written in compact form as:

$$\max_{\rho}\left\{\mathscr{S}[\rho]\,\middle|\,\mathscr{E}[\rho] = E,\ \forall k \in \mathbb{N}\quad \mathscr{Z}_k[\rho] = Z_k,\ \forall \mathbf{x} \in \mathscr{D}\quad \overline{1}(\mathbf{x})[\rho] = 1\right\}. \tag{100}$$

The probability measure ρ induced by the empirical density field has a concentration property. In other words, the average of an observable depending only on macrostates is dominated by the most probable macrostates, which are solutions of the variational problem (100).

An interesting limit case for the entropy in Eq. (89) is worth mentioning in order to relate the variational problem in Eq. (100) with previous studies on the shallow water system. Let us assume that there is neither height variation nor divergent fluctuations, which would be the case if considering a quasi-geostrophic model or 2d incompressible Euler equations. Then the only height level and the only divergence level are $\sigma_h = 1$, $\sigma_\mu = 0$, respectively. Defining $\rho_q(\mathbf{x}, \sigma_q) = \rho(\mathbf{x}, 1, \sigma_q, 0)$, Eq. (89) becomes up to an unimportant constant:

$$\mathscr{S}[\rho_q(\mathbf{x}, \sigma_q)] = -\int \mathrm{d}\mathbf{x}\mathrm{d}\sigma_q\ \rho_q \log \rho_q. \tag{101}$$

We recover in that case the macrostate entropy of the Miller-Robert-Sommeria theory [22,34].

Let us now assume that the height varies with position but that there is no local height fluctuations. Then at point $\mathbf{x}$ the only height level is $\sigma_h = \overline{h}(\mathbf{x}) = h(\mathbf{x})$. Defining $\rho_{q\mu}(\mathbf{x}, \sigma_q, \sigma_\mu) = \rho(\mathbf{x}, h(\mathbf{x}), \sigma_q, \sigma_\mu)$, the macrostate entropy in Eq. (89) becomes up to an unimportant constant:

$$\mathscr{S}\left[\rho_{q\mu}(\mathbf{x}, \sigma_q, \sigma_\mu), h(\mathbf{x})\right] = -\int \mathrm{d}\mathbf{x}\mathrm{d}\sigma_q\mathrm{d}\sigma_\mu\, h\rho_{q\mu} \log \rho_{q\mu}. \tag{102}$$

This form of the entropy was proposed by [7], without microscopic justification. Interestingly, [7] obtained Eq. (102) by assuming that the macrostate entropy can be written as

$$\mathscr{S} = \int \mathrm{d}\mathbf{x}\mathrm{d}\sigma_q\mathrm{d}\sigma_\mu h(\mathbf{x}) s\left(\rho_{q\mu}(\mathbf{x}, \sigma_q, \sigma_\mu)\right), \tag{103}$$

and by noting that $s(\rho) = -\rho \log \rho$ is the only function that leads to equilibrium states that are stationary.

4 General properties of the equilibria and simplification of the theory in limiting cases.

General properties of equilibrium states, solutions of the variational problem in Eq. (100), are discussed in this section. Critical points of the variational problem are given in the first subsection ; they are computed in appendix C. This allows to obtain an equation for the large scale flow, and to show that equilibrium states of the shallow water model are positive temperature states. A weak height fluctuation limit is considered in a second subsection. It is found that the large scale flow and the small scale fluctuations are decoupled in this limit, and that there is equipartition between small scale potential energy and small scale kinetic energy. The quasi-geostrophic limit is investigated in a third subsection.

4.1 Equilibrium states are stationary states with positive temperatures

4.1.1 Properties of the critical points

Critical points of the equilibrium variational problem defined in Eq. (100) are solutions of the equation

$$\forall \delta\rho, \quad \delta\mathscr{S}[\rho] - \beta\delta\mathscr{E}[\rho] - \sum_{k=0}^{+\infty}\alpha_k\delta\mathscr{Z}_k[\rho] - \int \mathrm{d}\mathbf{x}\,\xi(\mathbf{x})\int \mathrm{d}\sigma_h\mathrm{d}\sigma_q\mathrm{d}\sigma_\mu\,\delta\rho = 0 \qquad (104)$$

where $\beta, \{\alpha_k\}_{k\geq 0}$ and $\xi(\mathbf{x})$ are the Lagrange multipliers associated with the conservation of the energy, the potential vorticity moments, and with the normalization constraint, respectively. The computation of the critical points is performed in Appendix C.

A first key result of Appendix C is that solutions of Eq. (104) factorize as

$$\rho = \rho_h(\mathbf{x},\sigma_h)\,\rho_q(\mathbf{x},\sigma_q)\,\rho_\mu(\mathbf{x},\sigma_\mu)\,. \qquad (105)$$

A second result of Appendix C is that the pdf of the divergence field is a Gaussian

$$\rho_\mu(\mathbf{x},\sigma_\mu) = \sqrt{\frac{\beta}{2\pi}}\exp\left(-\frac{\beta}{2}\left(\sigma_\mu - \overline{\mu}\right)^2\right). \qquad (106)$$

Recalling that $\beta = \partial S/\partial E$ is the inverse temperature, we see that the variance of the pdf of the divergence field is given by the temperature of the flow:

$$\overline{\mu^2} - \overline{\mu}^2 = \beta^{-1}. \qquad (107)$$

This equation has important physical consequences. First, fluctuations of the divergent field do not vary with space. Second, the temperature of the equilibrium state is necessary positive. This contrasts with equilibrium states of two-dimensional Euler flow, which can be characterized by negative temperature states [33,23]. This was realized first by Onsager for the point vortex model [28]. In the context of 2D Euler equations in doubly periodic domains, the equilibrium states are always characterized by negative temperature, just as in the case discussed by Onsager (this result is shown in [2]). However, the existence of large scale flow structures characterized by positive temperatures are possible at low energy in the presence of lateral boundaries and non-zero circulation, and/or in the presence of bottom topography in the framework of QG model. Such positive temperature states are known and documented in the literature (see e.g. the original papers by [23], [34] and [18] or [4] and references therein).

In addition, it follows from Eqs. (95), (105) and (107) that the temperature is directly related to small scale kinetic energy due to the divergent velocity field:

$$\mathscr{E}_{\delta\mu} = \frac{1}{2\beta}. \qquad (108)$$

A third result of Appendix C is the expression of the pdf of the height field:

$$\rho_h(\mathbf{x},\sigma_h) = \frac{1}{\mathbb{G}_h(\mathbf{x})}\sigma_h^2\exp\left(-\beta\frac{g}{2}\sigma_h - \frac{\xi_h(\mathbf{x})}{\sigma_h}\right), \qquad \mathbb{G}_h(\mathbf{x}) = \int \mathrm{d}\sigma_h\,\sigma_h^2\exp\left(-\beta\frac{g}{2}\sigma_h - \frac{\xi_h(\mathbf{x})}{\sigma_h}\right), \qquad (109)$$

where $\xi_h(\mathbf{x})$ a function related to $\bar{h}(\mathbf{x})$ through $\bar{h} = \int \mathrm{d}\sigma_h\,\sigma_h\rho_h$.

A fourth result of Appendix C is the expression of the pdf of the potential vorticity field:

$$\rho_q(\mathbf{x},\sigma_q) = \frac{1}{\mathbb{G}_q(\beta\Psi_{mf})}\exp\left(\beta\Psi_{mf}\sigma_q - \sum_{k=1}^{+\infty}\alpha_k\sigma_q^k\right), \quad \mathbb{G}_q(\beta\Psi_{mf}) = \int \mathrm{d}\sigma_q\,\exp\left(\beta\Psi_{mf}\sigma_q - \sum_{k=1}^{+\infty}\alpha_k\sigma_q^k\right), \qquad (110)$$

where $\Psi_{mf}\left[\overline{h},\overline{q},\overline{\mu}\right]$ is the mass transport streamfunction of the mean flow defined through an Helmholtz decomposition of $\overline{h}\mathbf{u}_{mf}$:

$$\overline{h}\mathbf{u}_{mf} = \nabla^{\perp}\Psi_{mf} + \nabla\Phi_{mf}. \tag{111}$$

According to Eq. (110), the coarse-grained potential vorticity field is a function of the mass transport streamfunction:

$$\overline{q} = \frac{\mathbb{G}_q'(\beta\Psi_{mf})}{\mathbb{G}_q(\beta\Psi_{mf})}. \tag{112}$$

A fifth result of Appendix C is that the mass transport potential of the mean flow defined in Eq. (111) vanishes:

$$\Phi_{mf} = 0, \tag{113}$$

and the velocity field can now be written

$$\mathbf{u}_{mf} = \frac{1}{\overline{h}}\nabla^{\perp}\Psi_{mf}. \tag{114}$$

A sixth result of Appendix C concerns the Bernoulli function of the mean flow defined as

$$B_{mf} = \frac{1}{2}\mathbf{u}_{mf}^2 + g\left(\overline{h} + h_b - 1\right). \tag{115}$$

According to Eq. (218), combining Eq. (216) and Eq. (217) of Appendix C allows to express B_{mf} in terms of β, $\mathbb{G}_q$, $\mathbb{G}_h$ and a constant A_0 that can be computed in principle using the conservation of the total mass $\mathscr{Z}_0 = 1$:

$$B_{mf} = \beta^{-1}\log\left(\mathbb{G}_q\mathbb{G}_h\right) + g\overline{h} + A_0. \tag{116}$$

4.1.2 Equation for the large scale flow

Let us now establish the equations allowing to compute the large scale flow. A first equation is obtained by injecting the expression for the mean flow $\mathbf{u}_{mf}$ given in Eq. (114) into the expression of the Bernoulli function defined in Eq. (115), and by combining Eq. (115) with Eq. (116):

$$\frac{1}{2}\frac{1}{\overline{h}^2}\left(\nabla\Psi_{mf}\right)^2 + gh_b = \frac{1}{\beta}\log\left(\mathbb{G}_q\mathbb{G}_h\right) + A_1, \tag{117}$$

with $A_1 = A_0 + g$. A second equation is obtained by taking first the curl of Eq. (91), which yields $\overline{hq} - f = \nabla^{\perp}\mathbf{u}_{mf}$. Then, replacing $\mathbf{u}_{mf}$ by its expression given in Eq. (114), remembering that the potential vorticity field and the height field of the critical points of the variational problem are decorrelated ($\overline{qh} = \overline{h}\overline{q}$), and replacing $\overline{q}$ by its expression given in Eq. (112) yields

$$\overline{h}\frac{\mathbb{G}_q'\left(\beta\Psi_{mf}\right)}{\mathbb{G}_q\left(\beta\Psi_{mf}\right)} - f = \nabla\left(\frac{\nabla\Psi_{mf}}{\overline{h}}\right). \tag{118}$$

The closed system of partial differential equations (117) and (118) must be solved for $\overline{h}$ and Ψ_{mf} for a given value of β , $\mathbb{G}_q$, $\mathbb{G}_h$, and A_1. The set of parameters $\beta, \{\alpha_k\}_{k\geq 1}, A_1$ must *in fine* be expressed in terms of the constraints of the problem given by the energy E, and the potential vorticity moments $\{Z_k\}_{k\geq 0}$. This may require a numerical resolution in the general case.

We have seen in subsection 2.2 that a flow described by $(h,\mathbf{u})$ (or equivalently by (h,q,μ), or by (h,Φ,Ψ)) is a stationary state of the shallow water model if and only if $J(q,\Psi) = 0$ and $\Phi = 0$.

According to Eqs. (112) and (113), the large scale flow $\left(\overline{h},\mathbf{u}_{mf}\right)$ is a stationary state of the shallow water model.

4.1.3 Comparison with previous results

The set of equations (117) and (118) describing the large scale flow is similar to the one obtained by Weichman and Petrich [56] (through a Kac-Hubbard-Stratonovich transformation) and by Chavanis and Sommeria [7] (through a phenomenological generalization of the Miller-Robert-Sommeria theory), excepted that the rhs of Eq. (117) contains an additional term $\frac{1}{\beta}\log(\mathbb{G}_h)+A_1$ which does not appear in these previous works. The reason for the presence of this additional term is that we have taken into account the presence of small scale fluctuations of height and velocity, which were neglected in Refs. [56,7]. We will see in subsection 4.3 that this additional term becomes negligible with respect to the others in the quasi-geostrophic limit, in which case we recover exactly the set of equations for the large scale flow obtained in Refs. [56,7].

In addition, we have shown in subsection 4.1.1 that only positive temperature states are allowed, which shows that only one subclass of the states described by Eqs. (117) and (118) are actual equilibrium states.

4.2 Equipartition and decoupling in the limit of weak local height fluctuations

We consider in this subsection the limit of weak local height fluctuations. This step makes possible computation of explicit solutions of the variational problem (100). Meanwhile, it allows to explore the consequence of the presence of these small scale fluctuations on the structure of the large scale flow. By "limit of weak local height fluctuations", we mean

$$\forall \mathbf{x} \in \mathscr{D}, \quad \left(\overline{h^2}-\overline{h}^2\right)^{1/2} \ll \overline{h}. \tag{119}$$

As already argued in [56], this limit of weak local fluctuations is physically relevant, since the presence of shocks in the actual dynamics tends to dissipate small scale fluctuations of height or kinetic energy. This will be further discussed in subsection 5.2.

4.2.1 The height distribution

Assuming in addition that height levels σ_h such that $\left|\sigma_h-\overline{h}\right| \gg \left(\overline{h^2}-\overline{h}^2\right)^{1/2}$ do not contribute significantly to the pdf ρ_h defined in Eq. (109), we perform an asymptotic development with $\left(\sigma_h/\overline{h}-1\right) \ll 1$, and obtain at lowest order in this small parameter a Gaussian shape for the pdf:

$$\rho_h(\mathbf{x},\sigma_h) = \sqrt{\frac{g\beta}{2\pi\overline{h}(\mathbf{x})}}\exp\left(-\frac{g\beta}{2\overline{h}(\mathbf{x})}\left(\sigma_h-\overline{h}(\mathbf{x})\right)^2\right). \tag{120}$$

Similarly, the term $\mathbb{G}_h$ defined in Eq. (109) can be computed explicitly in this limit. One gets at lowest order

$$\mathbb{G}_h = \overline{h}^2(\mathbf{x})\sqrt{\frac{2\pi\overline{h}(\mathbf{x})}{g\beta}}\exp\left(-g\beta\overline{h}(\mathbf{x})\right). \tag{121}$$

Using Eq. (120), the weak local height fluctuation limit given in Eq. (119) can be interpreted as a low temperature limit $g\overline{h} \gg 1/\beta$. Injecting Eq. (121) in the previously established relation in Eq. (116) yields

$$B_{mf} = \frac{1}{\beta}\log\mathbb{G}_q + \frac{5}{2\beta}\log\overline{h} + A_2, \tag{122}$$

where A_2 is a free parameter determined by the conservation of the total volume. In the rhs of Eq. (122), there remains an additional term $\frac{5}{2\beta}\log\overline{h}$ which is not present in the large scale flow equation obtained by [56,7], but considering the weak flow limit allowed to obtain an explicit expression for this additional term. We will see in subsection 4.3 that this term becomes negligible with respect to $\frac{1}{\beta}\log \mathbb{G}_q$ in the quasi-geostrophic limit.

4.2.2 *Equipartition and the limit of weak small scale energy*

Injecting Eq. (120) in the expression of the small scale potential energy defined in Eq. (96) yields

$$\mathscr{E}_{\delta h} = \frac{1}{2\beta}. \tag{123}$$

Comparing this result with Eq. (108) shows equipartition of the small scale energy between the potential energy and the kinetic energy. The total energy due to small scale fluctuations is

$$E_{fluct} \equiv \mathscr{E}_{\delta h} + \mathscr{E}_{\delta \mu} = \frac{1}{\beta}. \tag{124}$$

This equipartition result of the small scale energy was already obtained by Warn when computing equilibrium states of the Galerking truncated dynamics in a weak flow limit [53]. More precisely, Warn decomposed the dynamics into vortical modes and inertia-gravity modes (which are defined as the eigenmodes of the linearized dynamics), and concluded that the energy of the equilibrium state should be equipartitioned among inertia-gravity modes in the limit of infinite wavenumber cut-off. This equipartition of energy among inertia-gravity modes would lead to equipartition between potential and kinetic energy at small scales, just as in our case. Here we have recovered this result with a different approach and we have generalized it beyond the weak flow limit.

Finally, we remark that the equipartition result of Eq. (124) shows that the low temperature limit $g\overline{h} \gg 1/\beta$ corresponds to a limit of weak small scale energy due to local fluctuation of the height field and of the divergent field:

$$g\overline{h} \gg E_{fluct}. \tag{125}$$

4.2.3 *Decoupling between the large scale flow and the fluctuations*

Still by considering the weak local height fluctuation limit, let us assume that E_{fluct} defined in Eq. (124) is given, which means that the temperature is given. According to Eq. (118) and Eq. (117), knowing the coarse-grained height field $\overline{h}$ and the pdf of potential vorticity levels $\rho_q(\mathbf{x}, \sigma_q)$ (which allows to compute $\mathbb{G}_q$) is sufficient to determine the mass transport streamfunction Ψ_{mf}, and hence the velocity $\mathbf{u}_{mf}$ by using Eq. (114). Then the mean-flow energy defined in Eq. (94) does not depend on ρ_μ:

$$\mathscr{E}_{mf}\left[\overline{h}, \rho_q\right] = \frac{1}{2}\int \mathrm{d}\mathbf{x}\, \left[\overline{h}\mathbf{u}_{mf}^2\left[\overline{h}, \overline{q}\right] + g\left(\overline{h} + h_b - 1\right)^2\right]. \tag{126}$$

The total energy defined in Eq. (97) is therefore the sum of the mean-field energy associated with a large scale flow that depends only on $\overline{h}$ and ρ_q , and of the energy of small scale fluctuations associated with height and divergence fluctuations:

$$\mathscr{E} = \mathscr{E}_{mf}\left[\overline{h}, \rho_q\right] + E_{fluct}. \tag{127}$$

In addition, injecting Eq. (105) in Eq. (92) gives the expression of the potential vorticity moments as a functional of $\overline{h}$ and ρ_q only:

$$\forall k \in \mathbb{N} \quad \mathscr{Z}_k\left[\overline{h}, \rho_q\right] = \int \mathrm{d}\mathbf{x}\mathrm{d}\sigma_q\, \overline{h}\rho_q\sigma_q^k. \tag{128}$$

Let us now consider the macrostate entropy functional defined in Eq. (89). Injecting Eqs. (106) and (120) in Eq. (105), and performing the asymptotic expansion of the integrand of the mean-field entropy with $\left(\sigma_h/\overline{h}-1\right) \ll 1$ leads at lowest order to (up to a irrelevant constant)

$$\mathscr{S} = \mathscr{S}_{mf}\left[\overline{h},\rho_q\right] + S_{fluct}\left(E_{fluc}\right), \tag{129}$$

$$\mathscr{S}_{mf}\left[\overline{h},\rho_q\right] \equiv -\int \mathrm{d}\mathbf{x}\mathrm{d}\sigma_q\, \overline{h}\rho_q \log \frac{\rho_q}{\overline{h}^{5/2}}, \tag{130}$$

$$S_{fluct}\left(E_{fluct}\right) \equiv \log\left(E_{fluct}\right). \tag{131}$$

Since ρ_q and $\overline{h}$ are two fields allowing to compute the large scale flow of energy $\mathscr{E}_{mf}$, and since the fluctuations of the potential vorticity field do not contribute to the small scale energy E_{fluct}, the entropy $\mathscr{S}_{mf}\left[\overline{h},\rho_q\right]$ will be referred to as the macrostate entropy of the large scale flow.

The second contribution to the total entropy in Eq. (129) depends only on the energy of the small scale fluctuations E_{fluct}. Since this energy is solely due to the local variance of the height field and of the divergent field, it will be referred to as the entropy of the small scale fluctuations. Note that in that case the height field is involved both in the large scale flow through its local mean value, and in the small scale fluctuations through its local variance.

The decoupling of the energy and the macrostate entropy functional into a part that depends only on $\rho_q,\overline{h}$ and another part that depends only on small scale height and divergent fluctuations with energy E_{fluct} has both a useful practical consequence and an interesting physical interpretation. The variational problem in Eq. (100) can now be recast into two simpler variational problems:

$$S\left(E,\{Z_k\}_{k\geq 0}\right) = \max_{E_{fluct}} \left\{ S_{mf}\left(E-E_{fluct},\{Z_k\}_{k\geq 0}\right) + S_{fluct}\left(E_{fluct}\right) \right\}, \tag{132}$$

$$S_{mf}\left(E_{mf},\{Z_k\}_{k\geq 0}\right) = \max_{\rho_q,\int\rho_q=1,\overline{h}} \left\{ \mathscr{S}_{mf}\left[\overline{h},\rho_q\right] \,\middle|\, \mathscr{E}_{mf}\left[\overline{h},\rho_q,\right] = E_{mf},\ \forall k\in\mathbb{N} \quad \mathscr{Z}_k\left[\rho_q,\overline{h}\right] = Z_k \right\}, \tag{133}$$

where $\mathscr{S}_{mf}$, $\mathscr{E}_{mf}$ and $\mathscr{Z}_k$ are the functional defined in Eqs. (130), (126), and (128), respectively. The variational problem in Eq. (132) describes two subsystems in thermal contact. In order to compute the equilibrium state, one can then compute independently the equilibrium state of each subsystem, and then equating their temperature in order to find the global equilibrium state. This classical argument follows directly from the maximization of (132). If the two subsystems can not have the same temperature (for instance when the temperature of both subsystem have a different sign), all the energy is stored in the subsystem with positive temperature in order to maximize the global entropy.

In the present case, a first subsystem is given by the large scale flow of energy E_{mf}, which involves the field $\overline{h}$ and the potential vorticity field described by the pdf of vorticity levels ρ_q. The equilibrium state of this subsystem is obtained by solving the variational problem in Eq. (133). This variational problem corresponds to the one introduced by [7] except that large scale flow energy is not the total energy but only the available energy when the energy of the fluctuations has been removed. The entropy of the large scale flow (130) in the variational problem (133) is closely related to the entropy introduced in [7] (up to a functional of $\overline{h}$). The potential vorticity moment constraints apply only to this subsystem. These additional constraints are essential since they allow for the possible existence of a large scale flow, see e.g. [4].

The second subsystem is given by the local small scale height fluctuations and the local small scale divergent fluctuations with total energy E_{fluct}, and with the entropy given in Eq. (131). The inverse temperature of this subsystem $\beta = \mathrm{d}S_{fluct}/\mathrm{d}E_{fluct}$, which, using Eq. (131), yields a temperature of $\beta^{-1} = E_{fluct}$.

In practice, it is easier to compute directly equilibrium states of the large scale flow subsystem in the canonical ensemble, where the energy constraint is relaxed. In that case the equilibrium states of this subsystem depend on the temperature E_{fluct} and on the dynamical invariants $\{Z_k\}_{k\geq 0}$. One just need to check a posteriori that this ensemble is equivalent to the microcanonical one by verifying that each admissible energy E_{mf} is reached when varying E_{fluct} form 0 to $+\infty$, for a given set of potential vorticity moments $\{Z_k\}_{k\geq 0}$. In order to find the actual equilibrium state associated with the total energy E, one then needs to solve the equation

$$E = E_{mf}\left(E_{fluct}, \{Z_k\}_{k\geq 0}\right) + E_{fluct}. \tag{134}$$

To conclude, it is now possible to study independently the large scale flow subsystem, to consider if necessary any approximation on this flow, such as the quasi-geostrophic limit or the Euler 2D limit, and finally to couple this subsystem with the small scale height and divergence fluctuations subsystem in order to select the actual equilibrium state. If one linearize the large scale flow entropy entropy (130) and the large scale flow energy (126), this picture of two subsystems in thermal contact gives a justification to the variational problem introduced by [45] and extended by [25]. [45, 25] suggest that for a given frozen in space potential vorticity, the dynamics should relax through geostrophic adjustment to a state minimizing the total energy. This result is recovered from the coupled variational problem (132).

4.2.4 Either a non-zero circulation or a non-zero bottom topography is required to sustain a large scale flow at equilibrium

It is shown in appendix D that when circulation is zero (i.e. when $Z_1 = f$) and when topography is zero ($h_b = 0$), a state with a large scale flow at rest ($\mathbf{u}_{mf} = 0$) is a maximizer of the large scale flow macrostate entropy among all the possible energies:

$$S(0, \{Z_k\}_{k\geq 0}) = \max_{E_{mf}} \left\{S_{mf}\left(E_{mf}, \{Z_k\}_{k\geq 0}\right)\right\} \quad \text{when } Z_1 = f \text{ and } h_b = 0. \tag{135}$$

According to Eq. (131), the fluctuation entropy S_{fluct} increases with the fluctuation energy E_{fluct}. The total macroscopic entropy $S_{mf}\left(E - E_{fluct}, \{Z_k\}_{k\geq 0}\right) + S_{fluct}\left(E_{fluct}\right)$ is therefore maximal when all the energy is transferred into fluctuations ($E_{fluct} = E$). This generalizes to a wider range of flow parameters and to a wider set of flow geometries a result previously obtained by Warn in a weak flow limit for a doubly periodic domain without bottom topography [53]. In addition, we will see in the next section that when there is a non-zero bottom topography and rotation, a large scale flow can be sustained at equilibrium.

4.3 The quasi-geostrophic limit

We show in this subsection that the variational problem of the Miller Robert Sommeria theory is recovered from Eq. (133) for the large scale flow subsystem when considering the quasi-geostrophic limit, which applies to strongly rotating and strongly stratified flows.

4.3.1 Geostrophic balance

Let $E_{mf}^{1/2}$ be the typical velocity of the large-scale flow and let $L = \sqrt{|\mathscr{D}|}$ be the typical horizontal scale of the domain where the flow takes place. We introduce the Rossby number and the Rossby radius of deformation respectively defined as

$$Ro \equiv \frac{E_{mf}^{1/2}}{fL}, \quad R \equiv \frac{\sqrt{gH}}{f}. \tag{136}$$

Here f is the Coriolis parameter, $H = 1$ is the mean depth and g the gravity, see subsection 2.1. If $f \neq 0$ we can always rescale time unit so that $f = 1$, and we make this choice in the following. It is also assumed that the aspect ratio of the domain where the flow takes place is of order one, so that $L = 1$ since we chose length unit so that $|\mathscr{D}| = 1$. The quasi-geostrophic limit corresponds to small Rossby number, and to a Rossby radius of deformation that is not significantly larger than the domain length scale:

$$Ro \ll 1, \quad R^{-1} = \mathscr{O}(1). \tag{137}$$

By construction, the mean flow is of the order of the Rossby number: $|\mathbf{u}_{mf}| \sim Ro$. The coarse-grained interface height is given by

$$\overline{\eta} = \overline{h} - 1 + h_b. \tag{138}$$

Let us assume that the spatial variations in fluid depth are small compared to the total depth $H = 1$, with the scaling $\overline{\eta} \sim R^{-2}Ro$. At lowest order in Ro, the mean flow Bernoulli potential defined in Eq. (115) becomes $B_{mf} = R^2\overline{\eta}$. Remembering that we consider in addition to the quasi-geostrophic limit a weak fluctuation limit given by Eq. (125), which can be expressed as $R^2 \gg \beta^{-1}$, Eq. (122) yields at lowest order

$$B_{mf} = R^2\overline{\eta} = \beta^{-1} \log \mathbb{G}_q + cst. \tag{139}$$

This equation implies $\mathrm{d}B_{mf}/\mathrm{d}\Psi_{mf} = \overline{q}$, consistently with what we expected for a stationary flow in the absence of small scale fluctuations, see subsection 2.2. Taking the curl of Eq. (139) and collecting the lowest order terms yields geostrophic balance[10]

$$R^2\nabla^{\perp}\overline{\eta} = \mathbf{u}_{mf}. \tag{140}$$

Eq. (140) also shows that the scaling hypothesis for $\overline{\eta}$ is self-consistent.

It is remarkable that equilibrium statistical mechanics predicts the emergence of geostrophic balance. We stress that those results, which have been obtained through the introduction of a semi-Lagrangian discrete model, are valid whatever the amplitude of bottom topography variations (i.e. beyond the usual approximation $h_b \sim Ro$ required to derive the quasi-geostrophic dynamics). By contrast, in the framework of a Eulerian discrete model, one would find that the large scale flow is not at geostrophic equilibrium unless $h_b \sim Ro$ or $h_b \ll Ro$, see Appendix E.

4.3.2 Quasi-geostrophic dynamics

The geostrophic balance is not a dynamical equation. When $h_b \sim Ro$, the dynamics if given by the quasi-geostrophic equations. At lowest order in Ro, we get $\overline{h}\mathbf{u}_{mf} = \mathbf{u}_{mf}$, and

$$\psi_{mf} = \Psi_{mf}, \quad \phi_{mf} = \Phi_{mf} = 0. \tag{141}$$

where Ψ_{mf} and ψ_{mf} are the transport streamfunction and the streamfunction obtained through the Helmholtz decomposition of $\overline{h}\mathbf{u}_{mf}$ and $\mathbf{u}_{mf}$, respectively. In that case, the relative vorticity is

$$\overline{\varpi} = \Delta\,\psi_{mf}. \tag{142}$$

The geostrophic balance (140) is equivalent to $\overline{\eta} = R^{-2}\psi_{mf} + C$. The value of ψ_{mf} at the domain boundary can always be chosen such that the constant term vanishes, which yields

$$\overline{\eta} = \frac{\psi_{mf}}{R^2}. \tag{143}$$

[10] For the shallow water model, the fluid is at hydrostatic balance. Thus the pressure in the fluid is $P(x,y,z,t) = P_o + \rho g(H + \eta(x,y,t) - z)$. Then the pressure horizontal gradient is simply proportional to the interface height horizontal gradient. Hence, the geostrophic balance simply writes $R^2\nabla^{\perp}\overline{\eta} = \mathbf{u}_{mf}$.

Mass conservation given in Eq. (2) leads then to the following constraint on the streamfunction:

$$\int \mathrm{d}\mathbf{x}\ \psi_{mf} = 0. \tag{144}$$

Given a potential vorticity level σ_q, a change of variable can be performed by introducing quasi-geostrophic potential vorticity levels[11]

$$\sigma_g \equiv (\sigma_q - 1)\,. \tag{145}$$

The pdf of quasi-geostrophic levels is

$$\rho_g(\mathbf{x}, \sigma_g) = \rho_q(\mathbf{x}, 1 + \sigma_g)\,. \tag{146}$$

The local quasi-geostrophic potential vorticity moments are defined as

$$\forall k \in \mathbb{N},\ \overline{q_g^k} \equiv \int \mathrm{d}\sigma_g\ \sigma_g^k \rho_g. \tag{147}$$

At lowest order in Ro, the coarse-grained quasi-geostrophic potential vorticity obtained by considering $k = 1$ in Eq. (147) becomes

$$\overline{q}_g = \overline{\omega} - \overline{\eta} + h_b. \tag{148}$$

which, using Eqs. (142) and (143), yields

$$\overline{q}_g = \Delta\psi_{mf} - \frac{\psi_{mf}}{R^2} + h_b\ . \tag{149}$$

In the quasi-geostrophic limit, the large scale flow is fully described by the streamfunction ψ_{mf}, which can be obtained by inverting the coarse-grained potential vorticity field defined in Eq. (149).

4.3.3 Quasi-geostrophic constraints and variational problem

At lowest order in Ro, and after an integration by part, the expression of the mean-flow energy in Eq. (126) is equal to the quasi-geostrophic energy:

$$\mathscr{E}_{mf,g}\left[\overline{q}_g\right] \equiv \frac{1}{2}\int \mathrm{d}\mathbf{x}\ \left[\left(\nabla\psi_{mf}\right)^2 + \frac{\psi_{mf}^2}{R^2}\right]\ , \tag{150}$$

where ψ_{mf} can be expressed in terms of $\overline{q}_g$ through Eq. (149). Similarly, the conservation of the potential vorticity moments defined in Eq. (128) implies the conservation of quasi-geostrophic potential vorticity moments

$$\forall k \in \mathbb{N}, \quad \mathscr{Z}_{g,k} \equiv \int \mathrm{d}\mathbf{x}\ \overline{q_g^k}\ , \tag{151}$$

and the macrostate entropy of the large scale flow defined in Eq. (130) writes now

$$\mathscr{S}_{mf,g}[\rho_g] = -\int \mathrm{d}\mathbf{x}\mathrm{d}\sigma_q\ \rho_g \log\rho_g. \tag{152}$$

[11] This change of variable is a guess guided by the fact that the QG potential vorticity is usually obtained by expanding the SW potential vorticity in the limit of small height variations minus a constant and unimportant term. Here we start by removing the unimportant constant (1 in our unit system) from the potential vorticity levels, and then perform the small scale expansion in height.

Using Eq. (150), (151) and (152), the variational problem in Eq. (133) is now recast into a simpler variational problem on the pdf ρ_g:

$$S_{mf,g}\left(E_{mf},\{Z_k\}_{k\geq 1}\right)=\max_{\rho_g,\int\rho_g=1}\left\{\mathscr{S}_{mf,g}\left[\rho_g\right]\middle|\,\mathscr{E}_{mf,g}\left[\rho_g\right]=E_{mf},\;\forall k\in\mathbb{N}\quad\mathscr{Z}_{g,k}\left[\rho_g\right]=Z_k\right\}. \tag{153}$$

The entropy $\mathscr{S}_{mf,g}$, the energy $\mathscr{E}_{mf,g}$ and the potential vorticity moments $\left\{\mathscr{Z}_{g,k}\right\}_{k\geq 1}$ are defined in Eqs. (152), (150) and (151). The variational problem defined in Eq. (153) is the variational problem of the Miller-Robert-Sommeria statistical mechanics [22,33].

4.3.4 Maximum energy states and consistency of the quasi-geostrophic approximation.

We have shown that in the weak height fluctuation limit and the quasi-geostrophic limit, computation of the large scale flow associated with the equilibrium state amounts to the computation of the solution of the variational problem in Eq. (153), with the restriction that the temperature is positive (due to the coupling with small scale fluctuations of height and divergence, as discussed in subsection 4.1.1). Here we discuss the solutions of this variational problem, which are energy maxima for a given set of potential vorticity moments $\left\{\mathscr{Z}_{g,k}\right\}_{k\geq 1}$. Although the initial weak fluctuation or quasi-geostrophic limit may not be fulfilled for such states, they can always be computed, and it provides an upper bound for the energy of the large scale flow obtained in those limits.

It is known that for positive temperatures ($\beta^{-1}=E_{fluct}>0$), the equilibrium entropy defined in Eq. (153) is concave [4], and the energy increases when β decreases. This means that the state with a maximum energy is reached when $\beta\to 0$, see e.g. [33].

Injecting $\beta=0$ in the expression of ρ_q in Eq. (110) leads to a uniform mean potential vorticity field. According to Eq. (146) and (147), this implies that the quasi-geostrophic potential vorticity field is also uniform: $\overline{q}_g=Z_{g,1}$ where $Z_{g,1}=\int \mathrm{d}\mathbf{x}\,\overline{q}_g$ is the circulation. Since $\overline{q}_g$ is a constant, this state is referred to as the “mixed” state. We get

$$\Delta\psi_{mix}-\frac{\psi_{mix}}{R^2}=Z_{g,1}-h_b. \tag{154}$$

Let us call E_{mix} the energy of the mixed state:

$$E_{mix}\equiv\max_{\rho_g,\int\rho_g=1}\left\{\mathscr{E}_{mf}|\forall k\in\mathbb{N}\quad\mathscr{Z}_{g,k}=Z_{g,k}\right\}. \tag{155}$$

Using Eq. (150), it can formally be written

$$E_{mix}=-\frac{1}{2}\int\left(Z_{g,1}-h_b\right)\left(\Delta-\frac{1}{R^2}\right)^{-1}\left(Z_{g,1}-h_b\right). \tag{156}$$

For a given domain geometry, a given circulation $Z_{g,1}$ and a given bottom topography field h_b, a non trivial large scale flow can be observed whenever $E_{mix}>0$. We see from Eq. (156) and (155) that the condition for a large scale flow to exist is that either the circulation $Z_{g,1}$ is non zero or the bottom topography h_b is non-zero. If both $Z_{g,1}=0$ and $h_b=0$, then the energy of the large scale flow vanishes ($E_{mf}=0$), and all the energy is lost in small scale fluctuations ($E=E_{fluct}$), consistently with the results of subsection 4.2.4. In this case, coupling thermally a large scale flow with fluctuations leads to a state with all the energy lost in fluctuations. Note that in the case $Z_{g,1}=0$ $R\sim 1$, E_{mix} can also be interpreted as a norm of the topography field h_b.

Since $E_{mf}\leq E_{mix,}$ and since E_{mix} depends only on the problem parameters (namely the circulation, the Rossby radius and the bottom topography), a sufficient condition to have $Ro\ll 1$ is

$$E_{mix}^{1/2}\ll 1. \tag{157}$$

If this condition is fulfilled, the quasi-geostrophic assumption is self-consistent (as well as for the scaling $h_b\sim Ro$ in the case $Z_{g,1}=0$).

5 Explicit computation of phase diagrams and discussion

The aim of this section is to apply the results of the previous section to the actual computation of equilibria and their energy partition. In order to solve analytically the variational problem of the statistical mechanics theory, we focus on a subclass of equilibria referred to as the energy-enstrophy equilibrium states. This allows to build phase diagrams in a two parameter space, and to discuss the energy partition between a large scale flow and small scale fluctuations when these parameters are varied. We finally discuss the role of shocks that occur in the actual shallow water dynamics, and present a geophysical application to the Zapiola anticyclone.

5.1 Energy-enstrophy equilibria for the quasi-geostrophic model.

We consider the variational problem

$$S_{g,mf}(E,Z_2) = \max_{\rho_g, \int \rho_g = 1} \left\{ \mathscr{S}_{mf,g}[\rho_g] \;\middle|\; \mathscr{E}_{mf,g}[\rho_g] = E_{mf},\ \mathscr{Z}_{g2}[\rho_g] = Z_2,\ \mathscr{Z}_{g1}[\rho_g] = 0 \right\}, \tag{158}$$

where the functionals $\mathscr{S}_{mf,g}$, $\mathscr{E}_{mf,g}$, $\mathscr{Z}_{g1,2}$ are defined in Eqs. (152), (150), and (151), respectively.

The peculiarity of this variational problem is that only two potential vorticity moments (the circulation and the enstrophy) have been retained as a constraint, in addition to the energy. Such energy-enstrophy equilibria are a subclass of statistical equilibria solutions of the more general variational problem given in Eq. (153), see e.g. [1,26]. For a given global distribution of potential vorticity, several limit cases on the energy allow to simplify the computation of the solutions of

the variational problem in Eq. (153) into the computation of the simpler variational problem in Eq. (158). For instance, assuming that bottom topography is non zero, and that the global potential distribution is such that the mixed state $\overline{q} = cst$ exists, the solutions of (158) are the solutions of the more general variational problem Eq. (153) when $E \to E_{mix}$.

The set of all potential vorticity fields $\overline{q}$ corresponding to solutions of the variational problem in Eq. (158) have been previously described by [6,49,51,26] in the case of a bounded geometry, and phase diagrams were obtained with energy E and circulation Z_1 as external parameters. The role of enstrophy Z_2 was not discussed. The main reason is that for a given large scale flow characterized by E_1 and Z_1, changing Z_2 would only imply changes in the small scale fluctuations of potential vorticity levels, see e.g. [26], and such small scale fluctuations do not contribute to the total energy. In the present case, Z_2 will play an important role in determining the energy partition between the large scale (vortical) flow and small scale fluctuations due to the height and divergent velocity field. For the sake of simplicity we consider vanishing circulations $Z_1 = 0$.

The problem (158) is solved in Appendix **F** for positive temperature states, and we present here the main results. A typical phase diagram is shown on Fig. 2. This figure is obtained by assuming that the bottom topography is proportional to the first Laplacian eigenmode (see Fig. 2-c) but it is explained in Appendix **F** that this phase diagram is generic to any bottom topography.[12]

There are two important quantities related to the height field: the maximum allowed energy E_{mix} defined in Eq. (156), and the available potential enstrophy

$$Z_b \equiv \int \mathrm{d}\mathbf{x}\, h_b^2, \tag{159}$$

which is the maximum reachable value for the macroscopic enstrophy $\int \mathrm{d}\mathbf{x}\, \overline{q_g}^2$, see Appendix **F**. Both Z_b and E_{mix} are a norm for the height field.

[12] For such a bottom topography, the topography, the stream function and the potential vorticity field are all proportional to each other for any initial condition for the enstrophy Z_2 and the energy E. That is why we do not show plots of the flows for different point of the phase diagram in Fig. 2. We rather choose to consider the case of the Zapiola drift in subsection 5.3 to see the effect of different value for the initial energy.

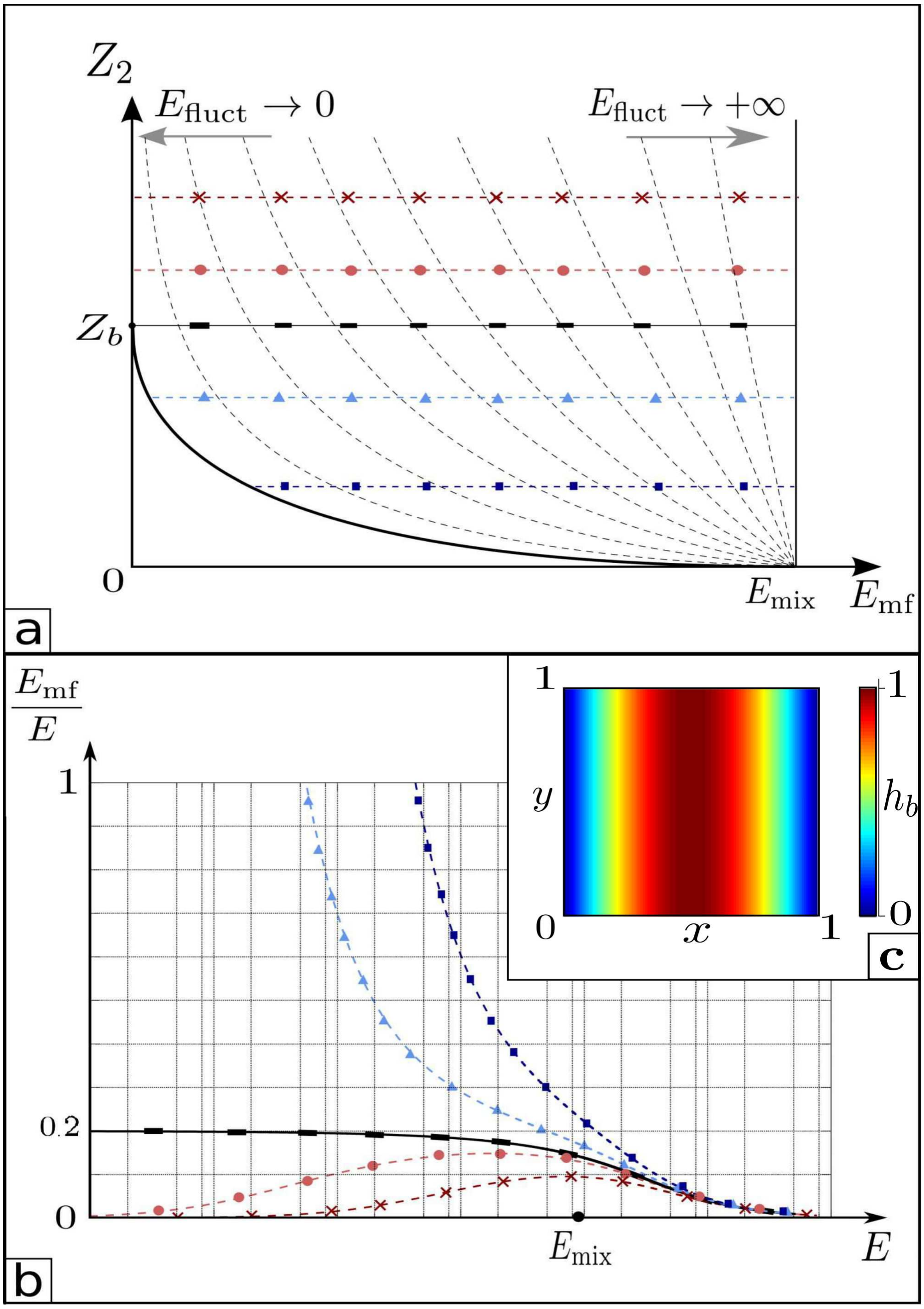

Fig. 2: a) Phase diagram of the energy-enstrophy ensemble in the plane (Z_2, E_{mf}). Z_b is the maximum reachable value for the macroscopic enstrophy defined in Eq. (159). E_{mix} is the maximum reachable energy for the mean-flow defined in Eq. (155). The dashed lines corresponds to isotherms, on which the energy E_{fluct} of the small scale fluctuations is constant. The thick black line on the bottom left corner is a boundary below which no equilibria exist. b) Ratio between the energy of the large scale flow E_{mf} over the total energy E_{mf}/E as a function of the total energy $E = E_{mf} + E_{fluct}$. The different curves correspond to different values of the initial microscopic enstrophy Z_2 represented by horizontal marked lines on panel a. The x-axis is on a logarithmic scale. c) Colormap plot of the bottom topography h_b used to compute the phase diagram in Fig. 2-a-b and 3. Here, $h_b = \sin(\pi x)$ is a single mode of the Laplacian operator. Thus the stream function ψ and the potential vorticity field q_g are simply proportional to h_b for any values of the initial enstrophy Z_2 and the initial energy E.

The phase diagram of the quasi-geostrophic energy-enstrophy ensemble restricted to positive temperature states is presented in Fig. 2-a. As explained in subsection 4.3, the energy of the large scale flow E_{mf} can not exceed the value E_{mix} defined in Eq. (155), due the restriction of positive temperature states. Depending on the sign of $Z_2 - Z_b$, where the potential enstrophy Z_b is defined in Eq. (159), the system behaves differently:

– When $Z_2 > Z_b$ the minimum admissible large scale flow energy is zero.
– When $Z_2 < Z_b$, there exists a minimum reachable large scale flow energy $E_{min}(Z_2)$ below which there is no equilibria. The curve E_{min} increases from 0 to E_{mix} when Z_2 decreases from Z_b to 0.

The thick black line in Fig. 2-a delimits the domain of existence for the equilibria. The thin dashed black curves represent the isotherm, i.e. the points of the diagram with the same value of E_{fluct}. Note that there is no bifurcation in this phase diagram; to each point (E_{mf}, Z_2) corresponds a single equilibrium state, whose expression is given explicitly in Appendix **F**. The structure of an equilibrium large scale flow above a topographic bump at low and high energy are presented in the last subsection.

The phase diagram in Fig. 2 allows to discuss the energy partition between small scale fluctuations and large scale flow when the quasi-geostrophic flow is coupled to small scale fluctuations of the height field and divergence field (through the shallow water dynamics). The motivation behind the works of [53] and [56] was the prediction of energy partition between large scale flow and small scale fluctuations. Here we provide for the first time an explicit expression for such energy partition.

We have seen that in the weak height fluctuation limit, the temperature of the large scale flow at equilibrium is given by the energy of the small scale fluctuations of height and divergence fields. The parameters are now the total energy E and enstrophy Z_2. The mean flow energy E_{mf} and the fluctuation energy E_{fluct} are found by solving

$$E = E_{mf}\left(E_{fluct}, Z_2\right) + E_{fluct}, \tag{160}$$

which is easily done graphically using the diagram of Fig. 2-a, and performed numerically in practice, see Appendix **F** for more details. The ratio of the large scale energy E_{mf} over the total initial energy E as a function of the total initial energy E is shown on Fig. 2-b for different values of initial enstrophy Z_2. According to computations performed in Appendix **F,** four cases for the energy partition in the low energy limit are distinguished depending on the sign of $Z_2 - Z_b$ and the scaling of $Z_2 - Z_b$ with E:

– When $Z_2 < Z_b$, the ratio E_{mf}/E tends to 1 when E tends to the minimal admissible energy $E_{min}(Z_2)$.
– When $Z_2 > Z_b$, the ratio E_{mf}/E tends to 0 when E tends to zero.
– When $Z_2 > Z_b$ with $Z_2 - Z_b \sim E^{\alpha}$ with $\alpha > 1/2$, the ratio E_{mf}/E tends to $1/5$ when E tends to zero.
– When $Z_2 > Z_b$ with $Z_2 - Z_b \sim E^{1/2}$, the ratio E_{mf}/E tends to a finite value (depending on the proportionality coefficient between $Z_2 - Z_b$ and $E^{1/2}$ and the bottom topography).
– When $Z_2 > Z_b$ with $Z_2 - Z_b \sim E^{\alpha}$ with $\alpha < 1/2$, the ratio E_{mf}/E tends to 0.

Whenever $Z_2 \geq Z_b$, we found $E_{fluct} \sim E$. We see on Fig. 2-b that the ratio E_{mf}/E converges to one when $E \to E_{min}(Z_2 < Z_b)$, that it converges to zero for $Z_2 > Z_b$, and that it converges to $1/5$ for $Z_b = Z_2$. Note also that E_{mf} is bounded by E_{mix} such that E_{mf}/E tends to zero when E tends to infinity.

5.2 The effect of energy dissipation and enstrophy dissipation

The actual shallow water dynamics is known to be characterized by shocks that prevent energy conservation. In addition, the presence of viscosity, no matter how small it is, may lead to enstrophy dissipation, and more generally would break the conservation of potential vorticity moments. The aim of this subsection is to discuss qualitatively the effect of these dissipative processes on the large scale flow, assuming that the system evolves through a sequence a equilibrium states, which is a natural hypothesis

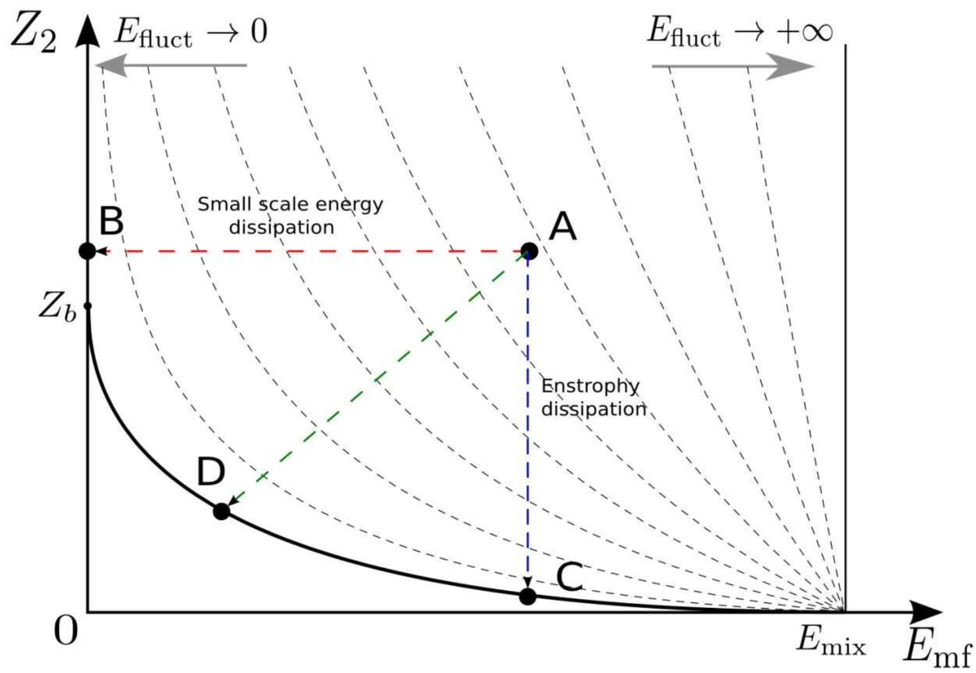

Fig. 3: Phase diagram of Fig. 2-a with hypothetical trajectories of the shallow water system in the presence of dissipation. Trajectory A-B: the dynamics dissipates small scale fluctuations of height and divergence field only. Trajectory A-C: the dynamics dissipates small scale fluctuations of potential vorticity. Trajectory A-D: the dynamics dissipates small scale fluctuations of all the fields.

if there exists a separation of time scales.

For the sake of simplicity, let us focus on the phase diagram obtained in the energy-enstrophy ensemble and described in the previous subsection. Let us consider that the system has reached at some time an arbitrary equilibrium state denoted by A in Fig. 3.

Let us first consider a case where energy is conserved, but enstrophy is dissipated. Then the system will evolve form point A to point C of Fig. 3. In other words, the small scale enstrophy Z_2 will be dissipated so that the enstrophy of the system tends to the minimum admissible value of enstrophy $Z_{2\,min}(E_{mf})$. Note that the curve $Z_{2\,min}(E_{mf})$ is nothing but the curve of energy minima for a given $Z_2 < Z_b$. We saw previously that the mean-flow energy E_{mf} dominates the fluctuation energy E_{fluc} on this line. We conclude that enstrophy dissipation alone drives the system towards a large scale flow without small scale fluctuations. This large scale flow vanished when topography vanished (since E_{mix} and Z_b would also vanish).

Let us now consider that the enstrophy is not dissipated and look the effect of energy dissipation. Let us first explain why one may expect that even weak dissipation can lead to a significant decrease of the large scale flow energy E_{mf} when $E_{fluct} \neq 0$ (in the absence of small scale fluctuations, E_{mf} would not decrease significantly with weak dissipation). At equilibrium, the energy of the small scale fluctuations E_{fluct} should be equipartitioned among all the modes of the height field and divergent field. Since there is an infinite number of such modes, this means a loss of energy through dissipative process, no matter how small they are. Since E_{mf} decreases with E_{fluct}, dissipating the energy E_{fluct} amounts to diminish the energy E_{mf}. For a given enstrophy $Z_2 > Z_b$, as in the case of point A Fig. (3), this dissipative process drives the system towards a zero energy state B. For a given enstrophy $Z_2 < Z_b$, this process would drive the system towards the line of minimal energy.

We expect to see both enstrophy and energy dissipation working together, so the trajectory of the system in phase diagram will be somewhere between the trajectory $A \to B$ and the trajectory $A \to C$, e.g. the trajectory e.g. $A \to D$. We conclude that in the presence of topography and small scale dissipation, we eventually reach a geostrophic regime at large time, provided that the initial enstrophy is sufficiently low (otherwise the final state contains no large scale flow). In addition, each time the system is perturbed by adding a little amount of energy without changing the enstrophy, it drives the system a bit more towards the mixed state $\left(E_{mf} = E_{mix}, Z_2 = 0\right)$.

5.3 Flow structure of the equilibrium states: application to the Zapiola Anticyclone.

So far we have discussed energy partition for shallow water equilibria in the quasi-geostrophic limit. Here we focus on the structure of the large scale flow associated with these equilibrium states. We consider for that purpose an oceanic application to the Zapiola anticyclone.

The Zapiola anticyclone is a strong anticyclonic recirculation taking place in the Argentine basin above a sedimentary bump known as the Zapiola drift [54,38,24]. The anticyclone is characterized by a mass transport as large as any other major oceanic current such as the Gulf Stream. It is a quasi-barotropic (depth independent) flow, with typical velocities of the order of 0.1 $m.s^{-1}$, and a lateral extension of the order of 800 km.

It is known from the earlier statistical mechanics studies that positive temperature states in the energy enstrophy ensemble of one layer models lead to anticyclonic circulations above topography anomalies, see e.g. [35] and references therein. A generalization of this result to the continuously stratified case with application to the Zapiola anticyclone is given in [48].

We have shown in this paper that quasi-geostrophic equilibria characterized by positive temperature states are also shallow water equilibria. The Zapiola anticyclone can therefore be interpreted as an equilibrium state of the shallow water model. Let us now show the qualitative difference between low energy states (when $Z_2 > Z_b$ and $E_{mf} \to 0$) and high energy states (when $E_{mf} \to E_{mix}$), which are shown on Fig. 4-a and Fig 4-b, respectively.

In both cases the bottom topography is the same, and its isolines are visualized with thin black lines. Bottom topography has been obtained from data available online and described in [39]. We isolated the largest close contour defining the Zapiola drift (the sedimentary bump above which the recirculation takes place), and considered the actual bottom topography inside this contour, and a flat bottom outside this contour. It allows to focus on the interesting flow structure occurring above the Zapiola drift.

Energy-enstrophy equilibrium states are then computed using Eq. (234) in Appendix F. We have seen that low energy limit corresponds to $\beta \to +\infty$ with $Z_2 > Z_b$. Taking this limit in Eq. (234) yields $\psi_{mf} = (Z_2 - Z_b)\, h_b/\beta$: the equilibrium state is a Fofonoff flow [12], meaning that streamlines are proportional to the isolines of topography, just as in Fig. 3-a.

The maximum energy state corresponds to the case $\beta = 0$, which corresponds to the mixed state defined in Eq. (154). The streamfunction is $\psi_{mix} = \left(R^{-2} - \Delta\right)^{-1} h_b$. The operator $\left(R^{-2} - \Delta\right)^{-1}$ with $R \sim 1$ is expected to smooth out the small scale topography features, just as in Fig. 4-a.

6 Conclusion

We have presented in this paper analytical computations of equilibrium states for the shallow water system, giving thus predictions for the energy partition into small scale fluctuations and large scale

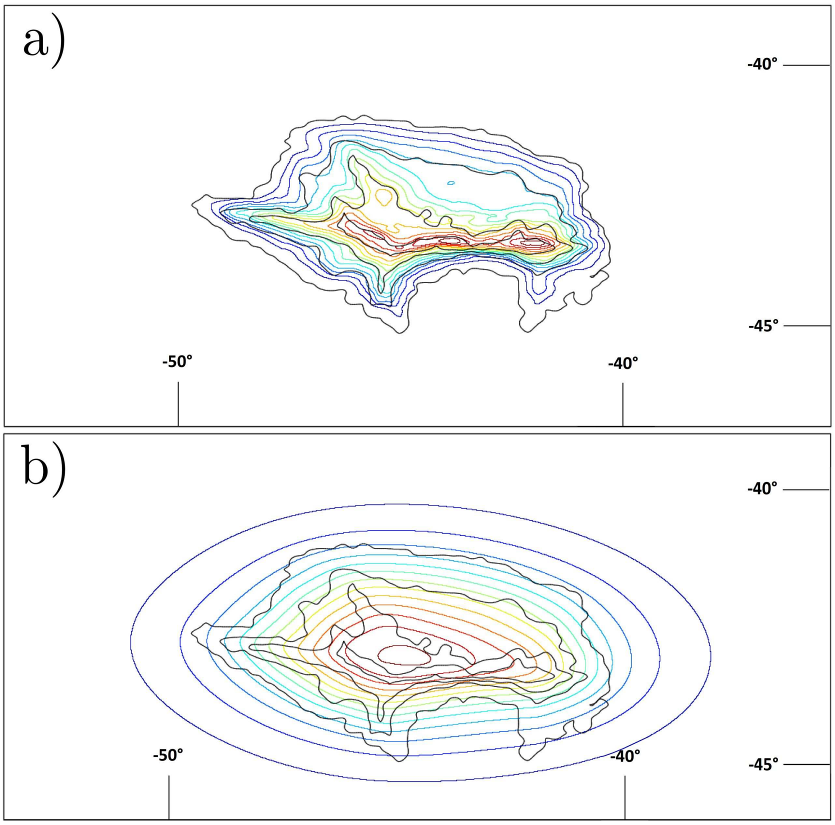

Fig. 4: Plots of the streamfunction isolines (colored lines) from higher values (red) to lower values (blue) over the Zapiola Drift topography iso-contours (black lines) for a small mean-flow energy (a)) and for a high mean-flow energy (b)).

flow. Our results rely on the definition of a discrete version of the shallow water model. Once our semi-Lagrangian discrete model was introduced, the whole machinery of equilibrium statistical mechanics could be applied.

We found that equilibrium states of the shallow water system are associated with the concomitant existence of a large scale flow which is a stationary state of the shallow water dynamics, superimposed with small scale fluctuations that may contain in some cases a substantial part of the total energy. The novelty of our work was to explicitly compute the contribution of these small scale fluctuations, and to decipher the physical consequences of the presence of these fluctuations.

In particular, we found that the presence of small scale fluctuations implies a positive temperature for the equilibrium state. This explains a previous result by Warn [53], who showed that equilibrium states in a weak flow limit admit no large scale flow when there is no bottom topography and no lateral boundaries. We have generalized these results, by showing that a large scale vortical flow exists at equilibrium when there is both rotation and bottom topography, or when there is a non-zero circulation.

In the limit of weak height fluctuations, we found equipartition of the small scale kinetic and potential energy. We also obtained an interesting physical picture of the equilibrium state, which may be interpreted in that limit as two subsystems in thermal contact. One subsystem is the “large scale” potential vortical flow whose entropy is closely related to the one introduced heuristically by Chavanis and Sommeria (see [7]). Our work provides therefore a microscopic justification of their entropy with a complete statistical mechanics derivation and a generalization of this results by including the presence of small scales fluctuations of the height and divergence fields. We note however that it is wrong to interpret the "large scale" potential vortical flow entropy as the entropy of the system, it is only one part of it. The other subsystem contains the field of height fluctuations associated with small scale potential energy, and the field of velocity fluctuations, associated with small scale kinetic energy. These velocity fluctuations are due solely to the divergent part of the velocity field. Warn obtained a similar result in the weak flow limit, by projecting the non-linear dynamics into eigenmodes of the linearized dynamics [53]. He found an energy partition into a vortical flow on the one side, and on inertia-gravity waves on the other side, with a weak coupling between both subsystems. Hence we may interpret local height and divergent velocity fields fluctuations appearing in our model as inertia gravity waves.

We studied the quasi-geostrophic limit for the large scale flow, taking into account the presence of small scale fluctuations. We recovered in this limit the variational problem of the Miller-Robert-Sommeria theory, with the additional constraint that the temperature is positive. We obtained phase diagrams in the particular case of energy-enstrophy equilibria, with explicit prediction for the ratio of energy between small scale fluctuations and a large scale quasi-geostrophic flow. This allowed to discuss the qualitative effect of small scale dissipation and shocks on the temporal evolution of the system. The main result is that such dissipative processes drive the system towards a minimum energy state (depending on the enstrophy), which may be non-zero.

In view of those results, the semi-Lagrangian discrete model seems to be a good discretization of the shallow water system. It has the key desired properties to take into account of the conservation of fluid particle volume, while working in an Eulerian framework. This gives a clear framework, which allows for a rigorous derivation once the discretization is assumed. Moreover we stress that it leads to equilibrium states that are stationary states of the shallow water model, by contrast with other choices of discretization. What is not completely satisfactory however is that there is a degree of arbitrariness in the definition of the discrete model. We have tried to have the model consistent with the geometric constraints related to the Liouville theorem, even though the link between the model and the Liouville theorem is clearly not rigorous. There is clearly room for improvement, but we are afraid we are faced with extremely tricky mathematical problems. Nevertheless we guess that the invariant measure of the discrete model converge to the actual invariant measure of the shallow water system in the continuous limit, but proving this is beyond the scope of this paper.

The statistical mechanics prediction of a vanishing large scale flow in the absence of boundary and bottom topography seems to contradict some numerical results performed in such configurations, in which case long lived vortex were reported, see e.g. [11]. This issue is related to the estimation of a time scale for the convergence towards equilibrium. Indeed, if the coupling between the large scale flow (potential vortical modes) and the small scale fluctuations (inertia-gravity waves) is weak, and if one starts from an initially balanced and unstable large scale flow, then this large scale flow may self-organize spontaneously on a short time scale into an equilibrium state of the quasi-geostrophic subsystem. This justifies the physical interest of the variational principle for the large scale flow introduced by [7] without small scale fluctuation. According to our statistical mechanics predictions, the energy of this large scale flow should leak into small scale fluctuations, but this process may be slow if coupling between both subsystems is weak. This difficult issue was already raised by Warn [53], and the interaction between geostrophic motion and inertia-gravity waves remains an active field of research [47]. In particular, very interesting models of interactions between near-inertial waves and geostrophic motion have been proposed [58, 14, 57]. In the context of statistical mechanics approaches, it has been proposed to compute equilibrium states with frozen degrees of freedom [36] or restricted partition functions in order to avoid the presence of inertia-gravity waves [16].

One of the main interests of the present study is the prediction of a concomitant large scale energy transfer associated with a large scale potential vortical flow and a small scale transfer of energy that is lost into small scale fluctuations interpreted as inertia-gravity waves. In that respect, the shallow water model lies between three dimensional and two-dimensional turbulence. There are other models for which the energy may be partitioned into a large scale flow and small scale fluctuations. There are, for instance, some strong analogies with the case of three-dimensional axisymmetric Euler equations, that is also intermediate between 2D and 3D flows. A statistical mechanics theory has been recently derived for this system by Thalabard et al [41]. This statistical theory can also be understood as two subsystems in contact, one of them being the fluctuations, and thus leading to positive temperatures [41]. From this key observation, [41] concluded that the temperature of the system, that measures the variance of the fluctuations, is positive and uniform in space, as previously stressed in [27]. More generally, and beyond those simplified flow models that allow for analytical treatment, it is common to observe in bounded laboratory experiment the emergence of large scale structures at the domain scale superimposed with small scale energy fluctuations, see e.g. [42]. It is not clear whether equilibrium theory may be relevant to describe such problems, but we believe that it is at least a useful first step to address these questions.

A Invariant measure and formal Liouville theorem

A.1 Formal Liouville theorem for the triplet of fields (h, hu, hv)

The existence of a formal Liouville theorem for the shallow water dynamics is shown in this appendix. The shallow water system is fully described by the triplet of fields (h, hu, hv). We consider a measure written formally as

$$\mathrm{d}\mu = C\mathscr{D}[h]\,\mathscr{D}[hu]\,\mathscr{D}[hv]\,,$$

with uniform density in (h, hu, hv)-space (C is a constant). The average of any functional A over this measure is

$$\forall\mathscr{A}[h, uh, vh]\,,\ \langle\mathscr{A}\rangle_\mu = \int \mathrm{d}\mu\ \mathscr{A}. \tag{161}$$

The term $\int \mathscr{D}[h]\,\mathscr{D}[hu]\,\mathscr{D}[hv]$ means that the integral is formally performed over each possible triplet of fields $(h,\ hu,\ hv)$. The measure is said to be invariant if

$$\forall\mathscr{A}\,,\quad \frac{\mathrm{d}}{\mathrm{d}t}\langle\mathscr{A}\rangle_\mu = 0 \tag{162}$$

This yields the condition

$$\forall\mathscr{A}\,,\quad \int \mathscr{D}[h]\,\mathscr{D}[hu]\,\mathscr{D}[hv]\int d\mathbf{x}\ \frac{\delta\mathscr{A}}{\delta h}\partial_t h + \frac{\delta\mathscr{A}}{\delta(hu)}\partial_t(hu) + \frac{\delta\mathscr{A}}{\delta(hv)}\partial_t(hu) = 0. \tag{163}$$

An integration by parts yields

$$\forall\mathscr{A}\,,\quad \int \mathscr{D}[h]\,\mathscr{D}[hu]\,\mathscr{D}[hv]\,\mathscr{A}\int d\mathbf{x}\ \left(\frac{\delta\partial_t h}{\delta h} + \frac{\delta\partial_t(hu)}{\delta(hu)} + \frac{\delta\partial_t(hv)}{\delta(hv)}\right) = 0. \tag{164}$$

We say that the equation follows a formal Liouville theorem if we formally have

$$\int d\mathbf{x}\ \left(\frac{\delta\partial_t h}{\delta h} + \frac{\delta\partial_t(hu)}{\delta(hu)} + \frac{\delta\partial_t(hv)}{\delta(hv)}\right) = 0, \tag{165}$$

which ensures that the measure $\mathrm{d}\mu$ is invariant.

The shallow water equations (5) and (6) can be written on the form

$$\partial_t(hu) = -\partial_x\left(hu^2 + \frac{1}{2}gh^2\right) - \partial_y(huv) + fhv, \tag{166}$$

$$\partial_t(hv) = -\partial_y\left(hv^2 + \frac{1}{2}gh^2\right) - \partial_x(huv) - fhu, \tag{167}$$

$$\partial_t h + \partial_x(hu) + \partial_y(hv) = 0. \tag{168}$$

We see that

$$\frac{\delta\partial_t h}{\delta h} + \frac{\delta\partial_t(hu)}{\delta(hu)} + \frac{\delta\partial_t(hv)}{\delta(hv)} = -\nabla\cdot\left(\frac{\delta(h\mathbf{u})}{\delta h} + \frac{\delta(hu\mathbf{u})}{\delta(hu)} + \frac{\delta(hv\mathbf{u})}{\delta(hv)}\right). \tag{169}$$

As the divergence operator is a linear operator, it commutes with the functional derivatives. This allows to conclude that the measure μ is invariant. This shows formally the existence of a Liouville theorem for the fields (h, uh, vh).

A.2 Change of variables from (h,hu,hv) to (h,q,μ)

The microcanonical measure can formally be written

$$\mathrm{d}\mu_{h,hu,hv} = \mathscr{D}[h]\,\mathscr{D}[hu]\,\mathscr{D}[hv]\,\delta(\mathscr{E}-E)\prod_{k=0}^{+\infty}\delta(\mathscr{Z}_k - Z_k)\,, \tag{170}$$

The constraints are more easily expressed in terms of the variables

$$h, \quad q = \frac{-\partial_y u + \partial_x v + f}{h}, \quad \mu = \Delta^{-1/2}(\partial_x u + \partial_y v)\,. \tag{171}$$

It will thus be more convenient to use these fields as independent variables. We call $J[(h,hu,hv)/(h,q,\mu)]$ the Jacobian of the transformation. We proceed step by step to compute this Jacobian. The change of variables $(h,hu,hv)\to(h,u,v)$ involves a upper-diagonal Jacobian matrix at each point $\mathbf{r}$:

$$J\left[\frac{(h,hu,hv)}{(h,u,v)}\right] = \begin{pmatrix} 1 & u & v \\ 0 & h & 0 \\ 0 & 0 & h \end{pmatrix} \tag{172}$$

which implies $\det(J[(h,hu,hv)/(h,u,v)]) = h^2$ and

$$\mathscr{D}[h]\,\mathscr{D}[hu]\,\mathscr{D}[hv] = h^2\mathscr{D}[h]\,\mathscr{D}[u]\,\mathscr{D}[v]\,. \tag{173}$$

The change of variable $(h,u,v)\to(h,\omega,\mu)$ involves linear operators that do not depend on space coordinates, thus the determinant of the Jacobian of the transformation is an unimportant constant:

$$\mathscr{D}[h]\,\mathscr{D}[u]\,\mathscr{D}[v] = C\mathscr{D}[h]\,\mathscr{D}[\omega]\,\mathscr{D}[\mu]\,. \tag{174}$$

Using $\omega = qh-1$, the change of variable $(h,\omega,\mu)\to(h,q,\mu)$ involves an upper diagonal Jacobian matrix at each point $\mathbf{r}$:

$$J\left[\frac{(h,\omega,\mu)}{(h,q,\mu)}\right] = \begin{pmatrix} 1 & q & 0 \\ 0 & h & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{175}$$

with a determinant $\det(J[(h,\omega,\mu)/(q,h,\mu)]) = h$. Finally, the Jacobian of the transformation is $J[(hu,hv,h)/(q,h,\mu)] = h^3$ and the microcanonical measure can formally be written

$$\mathrm{d}\mu_{h,q,\mu} = Ch^3\mathscr{D}[h]\,\mathscr{D}[q]\,\mathscr{D}[\mu]\,\delta(\mathscr{E}-E)\prod_{k=0}^{+\infty}\delta(\mathscr{Z}_k - Z_k)\,. \tag{176}$$

We note the presence of the pre-factor h^3 which gives the weight of each microscopic configuration in the (h,q,μ)-space.

B Relevance of the constraints for the discrete model

In this appendix, we explain how the dynamical invariants of the shallow water model, given in Eqs. (32) and (34) respectively, are related to the constraints of the microcanonical ensemble for the discrete model, given in Eqs. (92) and (97), respectively.

B.1 Areal coarse-graining for continuous fields

Let us consider a field $g(\mathbf{x})$ on the domain $\mathscr{D}$ where the flow takes place, and let us consider the uniform grid introduced in subsection 3.2.1. We define the local areal coarse-graining of the continuous field $g(\mathbf{x})$ over a site (i,j) as

$$\overline{g}_{ij} = N\int_{\text{site }(i,j)} \mathrm{d}\mathbf{x}\, g(\mathbf{x})\,, \tag{177}$$

where $1/N$ is the area of the site (i,j) and where $\int_{\text{site }(i,j)}$ means that we integrate over the site (i,j) only. With an abuse of notation, we use here the same notation $\overline{g}_{ij}$ as in Eq. (50), since the coarse-graining operator defined in Eq. (177)

generalizes to the continuous case the areal coarse-graining operator defined in Eq. (50) for the discrete microscopic model, taking into account the fact that for a fluid particle "n" of area $\mathrm{d}\mathbf{x}_n$ and height h_n, we get $N\mathrm{d}\mathbf{x}_n = 1/(Mh_n)$.

We denote $\overline{g}$ the continuous limit (large N) of $\overline{g}_{ij}$. Integrating a continuous field g amounts to perform the integration over its local average field $\overline{g}$:

$$\int \mathrm{d}\mathbf{x}\, g(\mathbf{x}) = \int \mathrm{d}\mathbf{x}\, \overline{g}(\mathbf{x}) \,. \tag{178}$$

B.2 Potential vorticity moments

Using (178), the potential vorticity moments in Eq. (34) simply leads to

$$\mathscr{Z}_k = \int \mathrm{d}\mathbf{x}\, \overline{hq^k}. \tag{179}$$

Now that the potential vorticity moments are expressed in terms of the areal coarse-graining of moments of h and q, it can directly be expressed in terms of the probability density field $\rho\left(\sigma_h, \sigma_q, \sigma_\mu\right)$ through Eq. (90), and we recover the expression of the constraint given in Eq. (92), whose discrete representation is given in Eq. (56).

B.3 Energy

Using (178), recalling that we restrict ourself to bottom topographies such that $\overline{h}_b = h_b$, the total energy of the shallow water model defined in Eq. (32) can be decomposed into a mean flow kinetic energy defined in Eq. (94), a potential energy term due to local height fluctuations and defined in Eq. (96), a fluctuating kinetic energy term

$$\mathscr{E}_{c,fluct} \equiv \mathscr{E} - \mathscr{E}_{mf} - \mathscr{E}_{\delta h} \tag{180}$$

$$\mathscr{E}_{c,fluct} \equiv \frac{1}{2}\int \mathrm{d}\mathbf{x}\, \left(\overline{h\mathbf{u}^2} - \overline{h}\mathbf{u}_{mf}^2\right), \tag{181}$$

where the velocity fields $\mathbf{u}$ and $\mathbf{u}_{mf}$ are computed from the triplet (h, q, μ) and from the triplet $\left(\overline{h}, \overline{hq}, \overline{h\mu}\right)$, respectively through

$$\mathbf{u} = \nabla^{\perp}\psi + \nabla\phi, \quad (hq - f) = \Delta\psi, \quad \mu = \Delta^{1/2}\phi, \tag{182}$$

and through

$$\mathbf{u}_{mf} = \nabla^{\perp}\psi_{mf} + \nabla\phi_{mf}, \quad \left(\overline{hq} - f\right) = \Delta\psi_{mf}, \quad \frac{\overline{h\mu}}{\overline{h}} = \Delta^{1/2}\phi_{mf}. \tag{183}$$

We want to discuss the relation between decomposition of the energy for the discrete model, Eq. (97) and the decomposition for the actual total energy defined in Eq. (32). Our construction is relevant if these two decomposition coincide in the continuous limit, or equivalently if $\mathscr{E}_{c,fluct}$ is equal to $\mathscr{E}_{\delta\mu}$ (95) in the continuous limit. In the following we show that this is the case if some cross correlations are actually negligible. More precisely, we assume that

1. For any positive integers k, l, m , the coarse-grained fields $\overline{h^k q^l \mu^m}(\mathbf{x})$ defined through the coarse graining procedure in Eq. (178) exist. In the framework of our microscopic model introduced in subsection 3.2, this hypothesis is automatically satisfied by assuming that the cut-off $\mu_{min}, \mu_{max}, q_{min}, q_{max}, h_{max}$ scales as N^{α} with $\alpha < 1$. Other fields may be characterized by local extreme values such that the limit defined in Eq. (178) does not converge. For instance, we will see that the actual divergence $\zeta = \Delta\phi$ of the equilibrium state is not bounded, i.e. that $\overline{\zeta}$ would have no meaning.
2. The fields h, q, μ are decorrelated (in particular, $\overline{h^k q^l \mu^m} = \overline{h^k q^l}\,\overline{\mu^m}$). This point will be shown to be self-consistent when computing the equilibrium state.
3. The coarse-grained divergent velocity field is equal to the mean-flow velocity field $\overline{\nabla\phi} = \nabla\phi_{mf}$.
4. $\overline{\Delta\left[\left(\phi - \phi_{mf}\right)^2\right]} = 0$. While $\nabla(\phi - \phi_{mf})$ is a random vector field characterized by wild local fluctuation, this hypothesis amounts to assume that those fluctuations have no preferential direction.

We believe that these four assumptions would be satisfied by a typical triplet of fields (h,q,μ) picked at random among all the possible states satisfying the constraints of the dynamics. By typical, we mean that an overwhelming number of fields would share these properties.

We then prove that these four assumptions are sufficient to prove that $\mathscr{E}_{c,fluct}$ is equal to $\mathscr{E}_{\delta\mu}$ (95) in the continuous limit. According to the assumption 1, $\overline{\varpi} = \overline{hq} - 1$ is well defined. Classical arguments show that the streamfunction of the coarse-grained vorticity field $\overline{\varpi}$ is equal to the streamfunction of the vorticity field, i.e. that$\nabla\psi_{mf} = \nabla\psi$, see e.g. [33,21]. Qualitatively, this is due to the fact that inverting the Laplacian operator smooth out local fluctuations of the vorticity[13]. This yields

$$\mathbf{u} = \mathbf{u}_{mf} + \nabla\left(\phi - \phi_{mf}\right). \tag{184}$$

Injecting this expression in the kinetic energy density expression Eq. (181), using that h and μ are not correlated (assumption 2), and using $\overline{\nabla\phi} = \nabla\phi_{mf}$ (assumption 3) yields

$$\mathscr{E}_{c,fluct} = \frac{1}{2}\int \mathrm{d}\mathbf{x}\,\overline{h}\,\overline{\left(\nabla\left(\phi - \phi_{mf}\right)\right)^2}. \tag{185}$$

Let us now remember the definition of the coarse-graining operator in Eq. (178):

$$\overline{\left(\nabla\left(\phi - \phi_{mf}\right)\right)^2} = \lim_{N\to\infty} N \int_{S_{ij}} \mathrm{d}\mathbf{x}\ \left(\nabla\left(\phi - \phi_{mf}\right)\right)^2, \quad \text{for } \mathbf{x} \in S_{ij}, \tag{186}$$

where S_{ij} is the surface covered by a grid site (i,j). An integration by parts yields

$$N\int_{S_{ij}} \mathrm{d}\mathbf{x}\ \left(\nabla\left(\phi - \phi_{mf}\right)\right)^2 = -N\int_{S_{ij}} \mathrm{d}\mathbf{x}\ \left[\left(\Delta^{-1/2}\left(\mu - \overline{\mu}\right)\right)\left(\Delta^{1/2}\left(\mu - \overline{\mu}\right)\right)\right] + N\oint_{\partial S_{ij}} \mathrm{d}l\ \mathbf{n}\cdot\left(\phi - \phi_{mf}\right)\nabla\left(\phi - \phi_{mf}\right) \tag{187}$$

Projecting the first term of the rhs on Laplacian eigenmodes allows to simplify the expression of the first term of the rhs in Eq. (187):

$$-N\int_{S_{ij}} \mathrm{d}\mathbf{x}\ \left[\left(\Delta^{-1/2}\left(\mu - \overline{\mu}\right)\right)\left(\Delta^{1/2}\left(\mu - \overline{\mu}\right)\right)\right] = N\int_{S_{ij}} \mathrm{d}\mathbf{x}\ \left(\mu - \overline{\mu}\right)^2. \tag{188}$$

The second term of the rhs in Eq. (187) can be written as

$$N\oint_{\partial S_{ij}} \mathrm{d}l\ \mathbf{n}\cdot\left(\phi - \phi_{mf}\right)\nabla\left(\phi - \phi_{mf}\right) = \frac{N}{2}\int_{S_{ij}} \Delta\left[\left(\phi - \phi_{mf}\right)^2\right] \tag{189}$$

which, according to assumption 4, vanishes in the large N limit. Finally, the kinetic energy density of the fluctuations is simply expressed as

$$\mathscr{E}_{c,fluct} = \frac{1}{2}\int \mathrm{d}\mathbf{x}\,\overline{h}\left(\overline{\mu^2} - \overline{\mu}^2\right). \tag{190}$$

Finally, we use again the assumption 2 to get

$$\mathscr{E}_{c,fluct} = \frac{1}{2}\int \mathrm{d}\mathbf{x}\ \left(\overline{h\mu^2} - \frac{\overline{h\mu}^2}{\overline{h}}\right) = \mathscr{E}_{\delta\mu}, \tag{191}$$

which is the expected result.

C Critical points of the mean-flow variational problem

In this Appendix, we compute the critical points of the mean-flow variational problem (100) stated in section 3. In a first step, we solve an intermediate variational problem in order to show the factorization of the probability density ρ with a Gaussian behavior for the divergence fluctuations. Knowing that, we solve in a second step the original variational problem.

[13] The divergent part of the velocity field can not be treated in the same way. Indeed, the operator $\Delta^{-1/2}$ is less smooth than the operator Δ^{-1}, and one can not derive $\phi = \phi_{mf}$ by inverting $\mu = \Delta^{1/2}\phi$. One may want to use $\zeta = \Delta\phi$, but the result would be the same since ζ is not bounded (the field μ is characterized by fluctuations which are controlled by the kinetic energy, and hence by the total energy, but this is not the case for ζ).

C.1 Intermediate variational problem

As the energy and the potential vorticity moments depend only on the coarse-grained fields $\overline{h}, \overline{h^2}, \overline{h\mu}, \overline{h\mu^2}$ and the local potential vorticity moments $\left\{\overline{hq^k}\right\}$, we introduce an intermediate variational problem where these coarse-grained fields are given as constraint:

$$\max_{\rho, \int\rho=1}\left\{\mathscr{S}[\rho]\;\middle|\;\overline{h},\overline{h^2},\overline{h\mu},\overline{h\mu^2},\left\{\overline{hq^k}\right\}_{k\geq 1}\text{ "fixed"}\right\}. \tag{192}$$

The idea of introducing the intermediate variational problem is to find a simpler ansatz for the probability density field ρ . This ansatz will be used afterward into the general variational problem (100).

In order to compute the critical points of the variational problem (192), we introduce the Lagrange multipliers $\alpha_h(\mathbf{x})$, $\alpha_{h2}(\mathbf{x})$, $\alpha_{h\mu}(\mathbf{x})$, $\alpha_{h\mu 2}(\mathbf{x})$, $\left\{\alpha_{hq,k}(\mathbf{x})\right\}_{k\geq 0}$ and $\xi(\mathbf{x})$ associated with the constraints $\overline{h}$, $\overline{h^2}$, $\overline{h\mu}$, $\overline{h\mu^2}$, $\left\{\overline{hq^k}\right\}_{k\geq 1}$ and the normalization constraint, respectively. Using Eq. (89) and the first variations

$$\forall\delta\rho,\ \delta S-\int\mathrm{d}\mathbf{x}\left[\alpha_h\delta\overline{h}+\alpha_{h2}\delta\overline{h^2}+\alpha_{h\mu}\delta\overline{h\mu}+\alpha_{h\mu 2}\delta\overline{h\mu^2}+\sum_{k=1}^{+\infty}\alpha_{hq,k}\delta\overline{hq^k}+\xi\int\delta\rho\right]=0, \tag{193}$$

leads to

$$\sigma_h\log\left(\frac{\rho}{\sigma_h^2}\right)+\sigma_h+\xi+\alpha_h\sigma_h+\alpha_{h2}\sigma_h^2+\alpha_{h\mu}\sigma_h\sigma_\mu+\alpha_{h\mu 2}\sigma_h\sigma_\mu^2+\sum_{k=1}^{+\infty}\alpha_{hq,k}\sigma_h\sigma_q^k=0 \tag{194}$$

We readily see from Eq. (194) that the probability density ρ factorizes into three decoupled probability densities ρ_q, ρ_h and ρ_μ corresponding respectively to the probability densities of the potential vorticity, the height and the divergence:

$$\rho=\rho_q(\mathbf{x},\sigma_q)\,\rho_h(\mathbf{x},\sigma_h)\,\rho_\mu(\mathbf{x},\sigma_\mu). \tag{195}$$

Using the constraints $\overline{\mu}=\int\mathrm{d}\sigma_\mu\ \sigma_\mu\rho_\mu$, $\overline{\mu^2}=\int\mathrm{d}\sigma_\mu\ \sigma_\mu^2\rho_\mu$, as well as the normalization constraints $\int\mathrm{d}\sigma_\mu\ \rho_\mu=\int\mathrm{d}\sigma_h\ \rho_h=\int\mathrm{d}\sigma_q\ \rho_q=1$, we get

$$\left\{\begin{array}{rcl}\rho_q(\mathbf{x},\sigma_q) & = & \dfrac{\exp\left(-\sum_{k=1}^{+\infty}\alpha_{hq,k}(\mathbf{x})\,\sigma_q^k\right)}{\int\mathrm{d}\sigma_q^{'}\exp\left(-\sum_{k=1}^{+\infty}\alpha_{hq,k}(\mathbf{x})\,\sigma_q^{'k}\right)}\\[2ex] \rho_h(\mathbf{x},\sigma_h) & = & \dfrac{\sigma_h^2\exp\left(-\alpha_{h2}(\mathbf{x})\,\sigma_h-\dfrac{\xi(\mathbf{x})}{\sigma_h}\right)}{\int\mathrm{d}\sigma_h^{'}\,\sigma_h^{'2}\exp\left(-\alpha_{h2}(\mathbf{x})\,\sigma_h^{'}-\dfrac{\xi(\mathbf{x})}{\sigma_h^{'}}\right)}\\[2ex] \rho_\mu(\mathbf{x},\sigma_\mu) & = & \dfrac{\exp\left(-\dfrac{1}{2}\dfrac{\left(\sigma_\mu-\overline{\mu}(\mathbf{x})\right)^2}{\overline{\mu^2}(\mathbf{x})-\overline{\mu}^2(\mathbf{x})}\right)}{(2\pi)^{1/2}\left(\overline{\mu^2}(\mathbf{x})-\overline{\mu}^2(\mathbf{x})\right)^{1/2}}\end{array}\right. . \tag{196}$$

We could now re-inject these expressions into the main variational problem (100), but only factorization and the Gaussian form of ρ_μ will be kept as an ansatz for ρ, which will simplify the computations. Thanks to this intermediate variational problem, we now know that the critical points of the original variational problem must be of the form:

$$\rho(\mathbf{x},\sigma_h,\sigma_q,\sigma_\mu)=\rho_h(\sigma_h,\mathbf{x})\,\rho_q(\sigma_q,\mathbf{x})\,\frac{\exp\left(-\frac{1}{2}\frac{(\sigma_\mu-\overline{\mu})^2}{\overline{\mu^2}-\overline{\mu}^2}\right)}{(2\pi)^{1/2}\left(\overline{\mu^2}-\overline{\mu}^2\right)^{1/2}}\ . \tag{197}$$

The entropy defined in Eq. (89) is therefore (up to a constant):

$$\mathscr{S}\left[\rho_h,\rho_q,\bar{\mu},\bar{\mu^2}-\bar{\mu}^2\right]=-\int\mathrm{d}\mathbf{x}\mathrm{d}\sigma_h\ \sigma_h\rho_h\log\left(\frac{\rho_h}{\sigma_h^2}\right)-\int\mathrm{d}\mathbf{x}\,\overline{h}\int\mathrm{d}\sigma_q\rho_q\log(\rho_q)+\int\mathrm{d}\mathbf{x}\,\frac{\overline{h}}{2}\log\left(\overline{\mu^2}-\overline{\mu}^2\right). \tag{198}$$

As a consequence of Eq. (197), the height field, the potential vorticity field and the divergence field are decorrelated. This property allows to rewrite the energy defined in Eq. (93)

$$\mathscr{E}\left[\rho_h,\rho_q,\overline{\mu},\overline{\mu^2}-\overline{\mu}^2\right]=\mathscr{E}_{mf}\left[\overline{h},\overline{q},\overline{\mu}\right]+\mathscr{E}_{\delta\mu}\left[\overline{h},\overline{\mu^2}-\overline{\mu}^2\right]+\mathscr{E}_{\delta h}\left[\overline{h},\overline{h^2}\right], \tag{199}$$

where

$$\begin{cases}\mathscr{E}_{mf}\left[\overline{h},\overline{q},\overline{\mu}\right]=\frac{1}{2}\int \mathrm{d}\mathbf{x}\left(\overline{h}\mathbf{u}_{mf}^2+g\left(\overline{h}+h_b-1\right)^2\right)\\ \mathscr{E}_{\delta\mu}\left[\overline{h},\overline{\mu^2}-\overline{\mu}^2\right]=\frac{1}{2}\int \mathrm{d}\mathbf{x}\,\overline{h}\left(\overline{\mu^2}-\overline{\mu}^2\right)\\ \mathscr{E}_{\delta h}\left[\overline{h},\overline{h^2}\right]=\frac{g}{2}\int \mathrm{d}\mathbf{x}\left(\overline{h^2}-\overline{h}^2\right)\end{cases} \tag{200}$$

with $\mathbf{u}_{mf}=\nabla^{\perp}\Delta^{-1}\left(\overline{q}\overline{h}-f\right)+\nabla\Delta^{-1/2}\overline{\mu}$. Similarly the potential vorticity moments (96) can be rewritten

$$\forall k \quad \mathscr{Z}_k\left[\overline{h},\overline{q^k}\right]=\int \mathrm{d}\mathbf{x}\,\overline{h}\overline{q^k} \tag{201}$$

where the coarse-grained moments are now defined as

$$\overline{h^l}=\int \mathrm{d}\sigma_h\,\sigma_h^l\rho_h, \qquad \overline{q^m}=\int \mathrm{d}\sigma_q\,\sigma_q^m\rho_q. \tag{202}$$

Thus, the general variational problem of the equilibrium theory given in Eq. (100) can be recast into a new variational problem on the independent variables $\rho_h(\mathbf{x},\sigma_h)$, $\rho_q(\mathbf{x},\sigma_q)$, $\overline{\mu}(\mathbf{x})$ and $\left[\overline{\mu^2}-\overline{\mu}^2\right](\mathbf{x})$:

$$S(E,D)=\max_{\substack{\rho_h,\rho_q,\overline{\mu},\overline{\mu^2}-\overline{\mu}^2\\ \int\rho_h=1,\int\rho_q=1}}\left\{\mathscr{S}\left[\rho_h,\rho_q,\overline{\mu},\overline{\mu^2}-\overline{\mu}^2\right]\;\middle|\;\mathscr{E}\left[\rho_h,\rho_q,\overline{\mu},\overline{\mu^2}-\overline{\mu}^2\right]=E,\ \forall k\quad \mathscr{Z}_k\left[\rho_h,\rho_q\right]=Z_k\right\}. \tag{203}$$

C.2 Computation of the critical points

In this subsection, we compute the critical points of the variational problem defined in Eq. (203). We introduce the Lagrange multiplier β, $\{\alpha_k\}_{k\geq 0}$, $\xi_q(\mathbf{r})$ and $\xi_h(\mathbf{r})$ associated respectively with the energy, the potential vorticity moments and the normalization constraints. Critical points of the variational problem (203) are solutions of

$$\forall\delta\rho_q,\delta\rho_h,\delta\overline{\mu},\delta\left(\overline{\mu^2}-\overline{\mu}^2\right),\quad \delta\mathscr{S}-\beta\delta\mathscr{E}-\sum_{k=0}^{+\infty}\alpha_k\delta\mathscr{Z}_k-\int\mathbf{dx}\left(\xi_q\int\delta\rho_q+\xi_h\int\delta\rho_h\right)=0. \tag{204}$$

The first variations of the macrostate entropy $\mathscr{S}$ (198) are

$$\begin{cases}\frac{\delta\mathscr{S}}{\delta\rho_h} &=-\sigma_h\left(\log\left(\frac{\rho_h}{\sigma_h^2}\right)+1\right)-\sigma_h\int \mathrm{d}\sigma_q\rho_q\log\left(\rho_q\right)+\sigma_h\frac{1}{2}\log\left(\overline{\mu^2}-\overline{\mu}^2\right)\\ \frac{\delta\mathscr{S}}{\delta\rho_q} &=-\overline{h}\left(\log\left(\rho_q\right)+1\right)\\ \frac{\delta\mathscr{S}}{\delta\overline{\mu}} &=0\\ \frac{\delta\mathscr{S}}{\delta\left(\overline{\mu^2}-\overline{\mu}^2\right)} &=\frac{\overline{h}}{2\left(\overline{\mu^2}-\overline{\mu}^2\right)}\end{cases}. \tag{205}$$

First variations of the energy given in Eqs. (199) and (200) contain three contributions: $\delta\mathscr{E}=\delta\mathscr{E}_{mf}+\delta\mathscr{E}_{\delta\mu}+\delta\mathscr{E}_{\delta h}$. The first contribution is

$$\delta\mathscr{E}_{mf}=\int \mathrm{d}\mathbf{x}\,\left[B_{mf}\delta\overline{h}+\left(\overline{h}\mathbf{u}_{mf}\right)\cdot\delta\mathbf{u}_{mf}\right], \tag{206}$$

where $B_{mf}=\mathbf{u}_{mf}^2/2+g\left(\overline{h}+h_b-1\right)$ is the mean-flow Bernoulli function defined in Eq. (115). Then, using the Helmholtz decompositions $\mathbf{u}_{mf}=\nabla^{\perp}\psi_{mf}+\nabla\phi_{mf}$ and recalling that $\overline{h}\mathbf{u}_{mf}=\nabla^{\perp}\Psi_{mf}+\nabla\Phi_{mf}$, two integrations by parts with the impermeability boundary condition yield

$$\delta\mathscr{E}_{mf}=\int \mathrm{d}\mathbf{x}\,\left[B_{mf}\delta\overline{h}-\Psi_{mf}\delta\Delta\psi_{mf}-\Phi_{mf}\delta\Delta\phi_{mf}\right]. \tag{207}$$

Using $\Delta\psi_{mf}=\overline{hq}-f$ and $\Delta^{1/2}\phi_{mf}=\overline{\mu}$ and the definition of the operator $\Delta^{1/2}$ leads to the final expression

$$\delta\mathscr{E}_{mf}=\int \mathrm{d}\mathbf{x}\ \left[\left(B_{mf}-\overline{q}\Psi_{mf}\right)\delta\overline{h}-\overline{h}\Psi_{mf}\delta\overline{q}-\Delta^{1/2}\Phi_{mf}\delta\overline{\mu}\right]. \tag{208}$$

Finally, we get:

$$\begin{cases}\frac{\delta\mathscr{E}}{\delta\rho_h} & =\sigma_h\left(B_{mf}-\overline{q}\Psi_{mf}+\frac{\overline{\mu^2}-\overline{\mu}^2}{2}+g\left(\frac{\sigma_h}{2}-\overline{h}\right)\right)\\ \frac{\delta\mathscr{E}}{\delta\rho_q} & =-\sigma_q\Psi_{mf}\overline{h}\\ \frac{\delta\mathscr{E}}{\delta\overline{\mu}} & =-\Delta^{1/2}\Phi_{mf}\\ \frac{\delta\mathscr{E}}{\delta\left(\overline{\mu^2}-\overline{\mu}^2\right)} & =\frac{\overline{h}}{2}\end{cases}, \tag{209}$$

First variations of the potential vorticity moments are

$$\forall k\in\mathbb{N}\quad \begin{cases}\frac{\delta\mathscr{Z}_k}{\delta\rho_h} & =\sigma_h\overline{q^k}\\ \frac{\delta\mathscr{Z}_k}{\delta\rho_q} & =\bar{h}\sigma_q^k\\ \frac{\delta\mathscr{Z}_k}{\delta\overline{\mu}} & =0\\ \frac{\delta\mathscr{Z}_k}{\delta\left(\overline{\mu^2}-\overline{\mu}^2\right)} & =0\end{cases}, \tag{210}$$

Injecting Eqs. (205), (209), and (210) in Eq. (204), and collecting the term in factor of $\delta\left(\overline{\mu^2}-\overline{\mu}^2\right)$ leads to

$$\overline{\mu^2}-\overline{\mu}^2=\frac{1}{\beta}. \tag{211}$$

Injecting Eq. (211) in the expression of ρ_μ given in Eq. (196) yields then

$$\rho_\mu\left(\mathbf{x},\sigma_\mu\right)=\left(\frac{\beta}{2\pi}\right)^{1/2}\exp\left(-\frac{1}{2}\beta\left(\sigma_\mu-\overline{\mu}\left(\mathbf{x}\right)\right)^2\right). \tag{212}$$

Similarly, collecting the term in factor of $\delta\rho_q$ in Eq. (204) leads to

$$0=\bar{h}\left(\log\left(p_q\right)+1\right)-\beta\sigma_q\Psi_{mf}\bar{h}+\sum_{k=0}^{+\infty}\alpha_k\bar{h}\sigma_q^k+\xi_q, \tag{213}$$

which, using the normalization constraint, leads to

$$\rho_q\left(\mathbf{x},\sigma_q\right)=\frac{1}{\mathbb{G}_q}\exp\left(\beta\sigma_q\Psi_{mf}-\sum_{k=1}^{+\infty}\alpha_k\sigma_q^k\right),\quad \mathbb{G}_q=\int \mathrm{d}\sigma_q^{,}\exp\left(\beta\sigma_q^{,}\Psi_{mf}-\sum_{k=1}^{+\infty}\alpha_k\sigma_q^{,k}\right). \tag{214}$$

Note that the sum inside the exponential is performed from $k=1$ to $k=+\infty$. The Lagrange parameter ξ_q has been determined using the normalization condition for the pdf.

Collecting the term in factor of $\delta\rho_h$ in Eq. (204) yields

$$-\left(\log\left(\frac{\rho_h}{\sigma_h^2}\right)+1\right)-\int \mathrm{d}\sigma_q\rho_q\log\left(\rho_q\right)-\frac{1}{2}\log\left(\beta\right)-\beta\left(B_{mf}+\frac{\beta^{-1}}{2}+g\left(\frac{\sigma_h}{2}-\overline{h}\right)\right)-\frac{\xi_h}{\sigma_h}+\beta\overline{q}\Psi_{mf}-\sum_{k\geq 0}\alpha_k\overline{q^k}=0, \tag{215}$$

which, using Eq. (214), leads to

$$-\left(\log\left(\frac{\rho_h}{\sigma_h^2}\right)+1\right)+\log\left(\mathbb{G}_q\right)-\frac{1}{2}\log\beta-\beta\left(B_{mf}+\frac{\beta^{-1}}{2}+g\left(\frac{\sigma_h}{2}-\overline{h}\right)\right)-\frac{\xi_h}{\sigma_h}-\alpha_0=0. \tag{216}$$

Using the fact that $\mathbb{G}_q$ and B_{mf} are fields depending only on $\mathbf{x}$, and using the normalization constraint for the pdf $\rho_h(\mathbf{x},\sigma_h)$, Eq. (216) yields

$$\rho_h\left(\mathbf{x},\sigma_h\right)=\frac{\sigma_h^2}{\mathbb{G}_h}\exp\left(-\beta\frac{g}{2}\sigma_h-\frac{\xi_h}{\sigma_h}\right),\quad \mathbb{G}_h=\int \mathrm{d}\sigma_h^{,}\,\sigma_h^2\exp\left(-\beta\frac{g}{2}\sigma_h^{,}+\frac{\xi_h}{\sigma_h^{,}}\right). \tag{217}$$

Injecting Eq. (217) back into Eq. (216) yields

$$B_{mf} = \beta^{-1}\log\left(\mathbb{G}_q\mathbb{G}_h\right) + g\overline{h} + \beta^{-1}\left(-\alpha_0 - \frac{3}{2} + \frac{1}{2}\log\beta\right). \tag{218}$$

One can notice that α_0 the Lagrange parameter related to the conservation of the total mass appears only here. Thus the last term $\beta^{-1}\left(\alpha_0 - 3/2 + \log\left(\beta\right)/2\right)$ in Eq. (218) can be computed from the conservation of the total mass $\mathscr{Z}_0 = Z_0$ and will be denoted A_0 in the following.

Collecting the terms in factor of $\delta\overline{\mu}$ in Eq. (204) leads to

$$\Phi_{mf} = 0. \tag{219}$$

D Global maximizers of the entropy of the large scale flow

We compute in this appendix an upper-bound for the macrostate entropy of the large scale flow defined in Eq. (130), for a given set of potential vorticity moment constraints defined in Eq. (92) (and arbitrary energy), and then show that when $Z_1 = f$ and $h_b = 0$, this upper bound for the macroscopic entropy is reached by the rest state.

This upper bound is the solution of the following variational problem:

$$S_{mf,max} = \max_{\substack{\overline{h},\rho_q \\ \int\rho_q=1}} \left\{\mathscr{S}_{mf}\left[\overline{h},\rho_q\right] \;\middle|\; \forall k \quad \mathscr{Z}_k\left[\overline{h},\rho_q\right] = Z_k\right\}. \tag{220}$$

Introducing Lagrange parameters $\{\gamma_k\}_{k\geq 0}$ associated with the potential vorticity moment constraints and the Lagrange parameter $\xi\left(\mathbf{x}\right)$ associated with the normalization constraint, the cancellation of first variations yields

$$\forall\delta\rho_q,\delta\overline{h}, \quad \delta\mathscr{S}_{mf} - \sum_{k=0}^{+\infty}\gamma_k\delta\mathscr{Z}_k + \int \mathrm{d}\mathbf{x}\;\xi\,\delta\overline{1} = 0. \tag{221}$$

The solution of this equation is

$$\rho_q = \frac{\exp^{-\sum_{k=1}^{+\infty}\gamma_k\sigma^k}}{\int\mathrm{d}\sigma\;\exp^{-\sum_{k=1}^{+\infty}\gamma_k\sigma^k}} \equiv \rho_{global}\left(\sigma\right) \tag{222}$$

where ρ_{global} depends only on the potential vorticity moments constraints $\{Z_k\}_{k\geq 1}$, and is independent from $\mathbf{x}$ and

$$\overline{h}\left(\mathbf{x}\right) = 1 \tag{223}$$

Note that the states characterized $\rho_q = \rho_{global}$, $\overline{h} = 1$ are solutions of the variational problem in Eq. (220), but this is only a subclass of the solutions of the variational problem of the equilibrium theory given in Eq. (153), which includes an additional energy constraint

We have shown in subsection (4.1) that for a given ρ_q, the large scale flow which is a solution (153) is obtained by solving Eqs. (117) and (118) for Ψ_{mf} and $\overline{h}$. Here we consider the particular case $\rho_q = \rho_{global}$ and $\overline{h} = 1$. One can compute $\overline{h}\overline{q}_{global} = \int\mathrm{d}\mathbf{x}\;\sigma\rho_{global} = Z_1$. We conclude that the large scale flow of the equilibrium state is also a global entropy maximizer, i.e. a solution of (220) when

$$Z_1 - f = \Delta\Psi_{mf}\;, \tag{224}$$

$$\frac{1}{2}\left(\nabla\Psi_{mf}\right)^2 + gh_b = A_2. \tag{225}$$

where $A_2 = \beta\log\mathbb{G}_q - A_1$ is a constant. We see that in the case ($Z_1 = f$, $h_b = 0$) , the solution of Eqs. (224) and (225) is the rest state $\Psi_{mf} = cst$ (with $A_2 = 0$). We conclude that the maximum of the macroscopic entropy of the large scale flow is reached by a flow at rest when there is no circulation ($Z_1 = f$) and no bottom topography ($h_b = 0$).

E Comparison with a Eulerian discrete model

The aim of this appendix is to discuss the construction of a possible invariant measure for the shallow water equations through an Eulerian discretization. We prove that the obtained equilibrium states differ from the one obtained through the semi-Lagrangian discretization used in the core of the paper. Moreover, we prove that the equilibrium states are not stationary states of the shallow water equations and that the statistical equilibria are not stable through coarse-graining.

We define a purely Eulerian discrete model by considering the same uniform $N\times N$ grid as for the semi-Lagrangian model, but assuming that each node is now divided into a finer $n\times n$ uniform microscopic grid. A microscopic configuration is given by the values of the fields (h,q,μ) for all the nodes of the microscopic grid:

$$y_{\text{micro}} \equiv \{h_{IJ,ij}, q_{IJ,ij}, \mu_{IJ,ij}\}_{\substack{1\le I,J\le N,\\ 1\le i,j\le n}} \tag{226}$$

where (I,J) and (i,j) correspond respectively to the position on the macroscopic grid and the position on the microscopic grid within the macroscopic node.

Contrary to the semi-Lagrangian model, the Eulerian model has the desired property to possibly be compatible with the formal Liouville theorem derived in Appendix A for the continuous dynamics (although no mathematical result exist). However, the volume of fluid varies from one microscopic grid node to another in the Eulerian model, depending on the value of the height $h_{IJ,ij}$. By comparison, our semi-Lagrangian approach respects the Lagrangian conservation laws (the height h is defined through the particle mass conservation). Because of the need to go through a discretization to build the microcanonical measure, we see that both the Eulerian and the semi-Lagrangian approaches necessarily break part of the geometric conservation laws of the continuous model. Hopefully rigorous mathematical proof of the convergence of the measures of one of the discretized model to an invariant measure of the continuous equations will settle rigorously this issue in a near future, however nobody seem to know how to attack this problem mathematically. We are thus led to the conclusion that based on current knowledge, there is no clear mathematical or theoretical *a priori* argument to choose either the Eulerian or the semi-Lagrangian discretization in order to guess the microcanonical measures. For now, the use of one discrete model or another to guess the microcanonical measure of the continuous shallow water equations can therefore only be justified *a posteriori*.

Let us now define the empirical density field as

$$d_{IJ}(\sigma_h,\sigma_q,\sigma_\mu)[y_{\text{micro}}] = \frac{1}{n}\sum_{i,j=1}^{n}\delta(h_{IJ,ij}-\sigma_h)\delta(q_{IJ,ij}-\sigma_q)\delta(\mu_{IJ,ij}-\sigma_\mu). \tag{227}$$

One can now compute the entropy of the macrostates $\rho = \{y_{micro}\,|\,\forall I,J \quad d_{IJ}\,[y_{micro}] = \rho_{IJ}\,\}$, which, after taking first the limit $n\to\infty$ and then the limit $N\to\infty$ leads to

$$\mathscr{S}_{\text{Eul}}[\rho] = -\int \mathrm{d}\mathbf{x}\mathrm{d}\sigma_h\mathrm{d}\sigma_q\mathrm{d}\sigma_\mu \quad \rho(\mathbf{x},\sigma_h,\sigma_q,\sigma_\mu)\log\left(\frac{\rho(\mathbf{x},\sigma_h,\sigma_q,\sigma_\mu)}{\sigma_h^3}\right). \tag{228}$$

This Eulerian macrostate entropy has to be compared with the macrostate entropy for the semi-Lagrangian discrete model given in Eq. (89). We can switch from expression to the other by changing ρ into $\sigma_h\rho$. We note that the two entropies become equivalent at lowest order in the limit of weak height fluctuations and weak height variations. However, in the general case, they are different, and therefore lead to different equilibrium states. In particular, is it straightforward to show that because of the absence of the factor σ_h in the expression of this Eulerian macrostate entropy (228), the critical points $\rho(\mathbf{x},\sigma_h,\sigma_q,\sigma_\mu)$ of the microcanonical variational problem do not factorize. Consequently, small scale height and velocity fluctuations of the equilibrium state are correlated. One can then show that those correlations are associated with non-zero Reynolds stresses in the momentum equations. In particular, the equilibrium state of Eulerian model satisfies

$$J(\Psi,\overline{q}) = -\overline{J(\Psi',q')} - \overline{J(\Phi',q')}, \tag{229}$$

where the r.h.s. is non-zero. If one removes those small scale fluctuations, the large scale flow is not a stationary state of the dynamics since $J(\Psi,\overline{q})\neq 0$. In other words, the equilibrium states of the Eulerian model are not stable by coarse-graining, contrary to the equilibria of the semi-Lagrangian model. Moreover, Eq. (229) and the properties of stationary states derived in subsection 2.2 imply that neither the potential vorticity field $\overline{q}$ nor the Bernoulli potential B_{mf} can simply be expressed as a function of Ψ_{mf}. As shown in subsection 4.3.1), the fact that B_{mf} is a function of Ψ_{mf} is essential to prove that the equilibrium is characterized by geostrophic balance at lowest order in the Rossby number Ro, when $Ro\to 0$. Consequently, the proof of geostrophic balance derived in the framework of the semi-Lagrangian model does not hold in the framework of the Eulerian model, unless the bottom topography is sufficiently small ($h_b \sim Ro$).

Let us finally argue that the stability by coarse-graining is a desirable physical property for the equilibrium states.

The first argument is a body of empirical observations. In either experiments, geophysical flows or numerical simulations flows governed by the shallow water equations (or the Navier-Stokes equations or the primitive equations in a shallow water regime) in the inertial limit (when they are subjected to weak forcing and dissipation, with a clear time scale separation) do actually self-organize and form large scale coherent structures for which there is a gradual decoupling of the flow large scales and small scales. A prominent example is the velocity field of Jupiter's troposphere.

The second argument follows. Macrostates that evolve through an autonomous equation, must increase the Boltzmann entropy. This is a general result in statistical mechanics, which is a consequence of the definition of the macrostate entropy as a Boltzmann entropy. Indeed as the Boltzmann's entropy measure the number of microstates corresponding to a given macrostate, it must increase for most of initial conditions. When there is furthermore a concentration property (which is the case for the shallow water case, both the the Eulerian and sem–Lagrangian discretizations), the number of initial conditions for which the entropy can decrease decays exponentially with N (N is often the number of particles in statistical mechanics, here the number of degrees of freedom of our discretization). As a consequence, the set of equilibrium macrostates (entropy maxima) has to be stable through the dynamics for most initial conditions. In the shallow water case, in statistical equilibrium, obtained either using the semi-Lagrangian or Eulerian discretization discussed above, the stream function concentrates close to a single field (the stream function fluctuations vanish in the large N limit). As a consequence the macrostate stream function, which is a single field thanks to this concentration property, has to be stationary for the dynamics. Those two properties, that follow from the definition of the Boltzmann entropy, are actually verified for the equilibrium measure constructed from a semi-Lagrangian discretization, but not for the equilibrium measure constructed from a Eulerian discretization. For this reason, we conclude that the microcanonical measure constructed from the purely Eulerian discretization is inconsistent with the shallow water dynamics.

We note moreover that the stability of the equilibrium macrostates through coarse graining ensures that the equilibrium states of the inviscid system are not affected by perturbations such as a weak small scale dissipation in momentum equations. This property is not a-priori required for the invariant measure of the shallow-water equations. However It is extremely interesting as it is a hint that this invariant measure may be relevant for non perfect flow in the inertial limit.

F Energy-Enstrophy ensemble

F.1 Computation of the critical points

In this Appendix, we compute the solutions of the variational problem (158) and describe the corresponding phase diagram. Critical points of the variational problem (158) are computed through the variational principle:

$$\forall \delta\rho_g, \quad \delta\mathscr{S}_{mf,g} - \frac{1}{E_{fluct}}\delta\mathscr{E}_{mf,g} - \gamma_2\delta\mathscr{Z}_{g2} - \gamma_1\delta\mathscr{Z}_{g1} - \int \mathrm{d}\mathbf{x}\,\xi(\mathbf{x})\int \mathrm{d}\sigma_q\,\delta\rho_g = 0, \tag{230}$$

where γ_2, γ_1 and $\xi(\mathbf{x}$ are Lagrange multipliers associated with the enstrophy conservation, the circulation conservation, and the normalization respectively. Anticipating the coupling between the large scale quasi-geostrophic flow and the small scale fluctuations, the temperature is denoted E_{fluct} (the inverse temperature is the Lagrange parameter associated with energy conservation).This yields

$$\rho_g(\mathbf{x},\sigma_q) = \sqrt{\frac{1}{2\pi\left(Z_2-\overline{Z}_2\right)}}\exp\left[-\frac{1}{2\left(Z_2-\overline{Z}_2\right)}\left(\sigma_q - \left(\frac{\psi_{mf}}{E_{fluct}}-\gamma_1\right)\left(Z_2-\overline{Z}_2\right)\right)^2\right] \tag{231}$$

where we have introduced the enstrophy of the coarse-grained potential vorticity

$$\overline{Z}_2 \equiv \int \mathrm{d}\mathbf{x}\overline{q}_g^2. \tag{232}$$

Injecting (231) in Eq. (147) , using the mass conservation constraint given in Eq. (144) and the zero circulation constraint $\mathscr{Z}_1\left[\overline{q}_g\right] = 0$ yields

$$\overline{q}_g = \widetilde{\beta}\psi_{mf} \quad \text{with } \widetilde{\beta} \equiv \frac{\left(Z_2-\overline{Z_2}\right)}{E_{fluct}}. \tag{233}$$

Note that $\widetilde{\beta}$ is necessarily positive given that $Z_2 - \overline{Z_2} \geq 0$. Injecting Eq. (233) in Eq. (149), the streamfunction can be computed explicitly by solving

$$\widetilde{\beta}\psi_{mf} = \Delta\psi_{mf} - \frac{1}{R^2}\psi_{mf} + h_b. \tag{234}$$

In order to solve this equation, it is convenient to introduce the Laplacian eigenmodes of the domain $\mathscr{D}$, with $k \in \mathbb{N}^+$:

$$\Delta e_k = -\lambda_k^2 e_k \quad \text{with } e_k = 0 \text{ on } \partial\mathscr{D}, \tag{235}$$

where the eigenvalues $-\lambda_k^2$ are arranged in decreasing order. We assume those eigenvalues are pairwise distinct. We also assume that the bottom topography is sufficiently smooth such that $\sum_k |h_{bk}|^2 \lambda_k^2 < +\infty$. Then, given that $\widetilde{\beta} > 0$, the projection of the mean flow streamfunction on the Laplacian eigenmode $e_k(\mathbf{x})$ is obtained directly from Eq. (234):

$$\psi_k = \frac{h_{bk}}{\widetilde{\beta} + \lambda_k^2 + R^{-2}}. \tag{236}$$

We see that there is a unique solution ψ_{mf} for each value of $\widetilde{\beta}$. This solution is therefore the equilibrium state. All the large scale flows associated with statistical equilibrium states of the shallow water system restricted to the energy-enstrophy ensemble with zero circulation are obtained from Eq. (236) when varying $\widetilde{\beta}$ from 0 to $+\infty$.

F.2 Construction of the phase diagram

The problem is now to find which equilibrium state is associated with the constraints (E, Z_2). In the following, we explain how to find the equilibrium states associated with parameters (E_{mf}, Z_2), and how to compute the temperature E_{fluc} for each of those states. It is then straightforward to obtained the total energy $E = E_{mf} + E_{fluc}$.

Injecting Eq. (236) in the expression of the quasi-geostrophic mean-flow energy defined in Eq. (150) yields

$$E_{mf} = \frac{1}{2} \sum_{k=1}^{+\infty} \left(\lambda_k^2 + R^{-2}\right) \left(\frac{|h_{bk}|}{\widetilde{\beta} + \lambda_k^2 + R^{-2}} \right)^2. \tag{237}$$

The mixing energy E_{mix} defined in Eq. (155) is recovered for $\widetilde{\beta} = 0$, given that $Z_1 = 0$. In the range $\widetilde{\beta} > 0$, the energy E_{mf} is a decreasing function of $\widetilde{\beta}$, varying from $E_{mf} = E_{mix}$ to $E_{mf} = 0$, see Fig. 5-b and Fig. 5-d.

Injecting Eq. (236) in the expression of the macroscopic enstrophy given in Eq (232) yields

$$\overline{Z}_2 = \sum_{k=1}^{+\infty} |h_{bk}|^2 \left(1 - \frac{\lambda_k^2 + R^{-2}}{\widetilde{\beta} + \lambda_k^2 + R^{-2}} \right)^2. \tag{238}$$

The potential enstrophy Z_b defined in Eq. (159) is recovered for $\widetilde{\beta} = +\infty$. The macroscopic enstrophy $\overline{Z}_2$ is an increasing function of $\widetilde{\beta}$, varying from $\overline{Z_2} = 0$ (for $\widetilde{\beta} = 0$) to $\overline{Z_2} = Z_b$ for ($\widetilde{\beta} = +\infty$), see Fig. 5-a and Fig. 5-c.

Two expressions of the macroscopic enstrophy $\overline{Z_2}$ in terms of the parameters $\widetilde{\beta}$ have been obtained: one is given by Eq. (238), the other arises from the definition of $\widetilde{\beta}$ in Eq. (233), which yields

$$\overline{Z_2} = Z_2 - E_{fluct}\widetilde{\beta}. \tag{239}$$

For given values of E_{fluct} and Z_2 , the values of $\widetilde{\beta}$ and $\overline{Z_2}$ are obtained by finding the intersection between the two curves defined in Eq. (238) and (239), respectively. Once $\widetilde{\beta}$ is obtained, Eq. (237) gives directly the value of the mean-flow energy E_{mf}. The phase diagram presented in Fig. 2 is obtained numerically by using this procedure. Graphical arguments presented in the following allow to understand the structure of this phase diagram.

F.3 Limit cases for the energy partition

Let us first note through Figs. 5-a and 5-c that $\widetilde{\beta}$ is an decreasing function of E_{fluct}. Indeed, $\widetilde{\beta}$ is given by the intersection between the solid curve representing the expression of $\overline{Z}_2$ given Eq. (238) and the dashed line representing the affine expression of $\overline{Z}_2$ given Eq. (239) where $-E_{fluct}$ is the slope. Then we know that the total energy $E = E_{mf}(\widetilde{\beta}) + E_{fluct}$ is an increasing function of E_{fluct}. Let us now consider different limit cases.

The limit $\mathbf{E} \to \infty$ with $\mathbf{Z_2}$ fixed: In this limit, we have also $E_{fluct} \to \infty$. and $\widetilde{\beta} \to 0$ Hence, one gets from Eq. (237) (see also Figs. 5-b and 5-d):

$$\lim_{E \to +\infty} E_{mf} = E_{mix}, \quad \lim_{E \to +\infty} \frac{E_{mf}}{E} = 0. \tag{240}$$

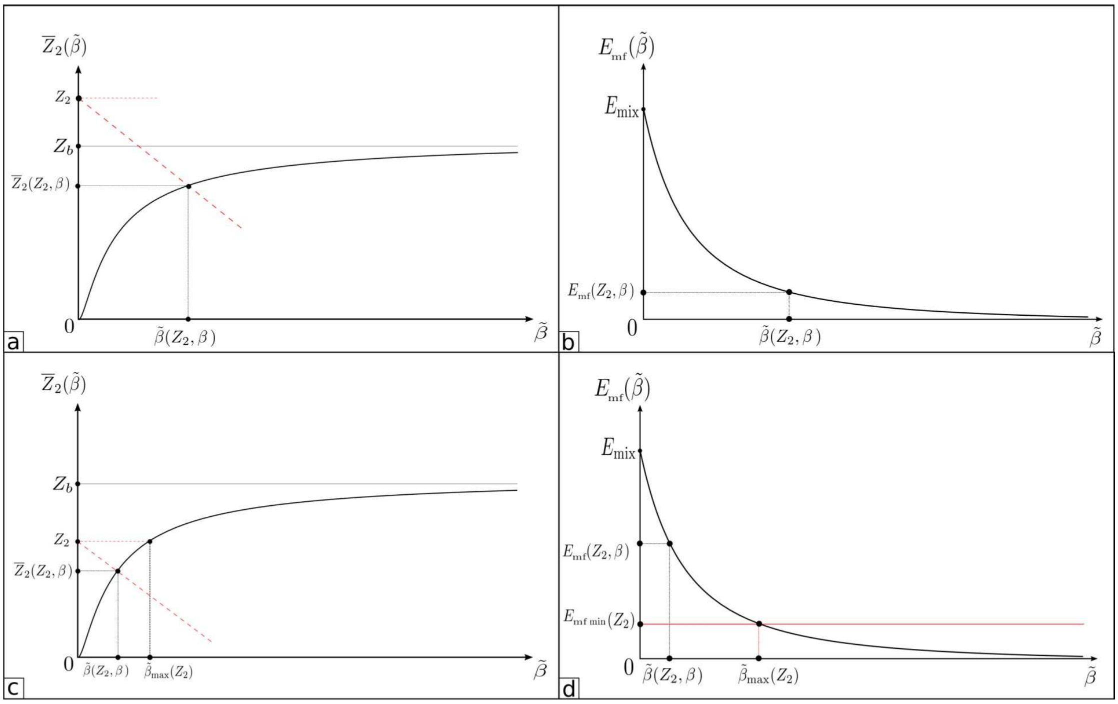

Fig. 5: a) Variation of the macroscopic enstrophy Z_2 over $\widetilde{\beta}$, case $Z_2 > Z_b$. b) Variation of the mean-flow energy E_{mf} with $\widetilde{\beta}$, case $Z_2 < Z_b$. c) Variation of the macroscopic enstrophy Z_2 with $\widetilde{\beta}$, case $Z_2 < Z_b$. d) Variation of the mean-flow energy E_{mf} with $\widetilde{\beta}$, case $Z_2 < Z_b$.

The lowest E limit with $\mathbf{Z_2 < Z_b}$ fixed: In this limit, we have also $E_{fluct} \to 0$. One gets from Fig. 5-c that $\widetilde{\beta} \to \widetilde{\beta}_{max}(Z_2)$. Hence, E_{mf} reaches a minimum admissible energy $E_{min}(Z_2) = E_{mf}\left(\widetilde{\beta}_{max}(Z_2)\right)$. Then:

$$\lim_{E \to E_{min}(Z_2)} \frac{E_{mf}}{E} = 1. \tag{241}$$

The limit $\mathbf{E \to 0}$ with $\mathbf{Z_2 > Z_b}$ fixed: In this limit, we have also $E_{fluct} \to 0$. One gets from Fig. 5-a that $\widetilde{\beta} \to \infty$ and that $\overline{Z}_2 \to Z_b$. Hence, from Eqs. (237) and (239), we obtain:

$$\begin{cases} E_{mf} & = C_b \widetilde{\beta}^{-2} + o\left(\widetilde{\beta}^{-2}\right) \\ \widetilde{\beta} & \underset{E \to 0}{\sim} (Z_2 - Z_b) E_{fluct}^{-1} \end{cases}, \quad \text{with} \quad C_b = \frac{1}{2} \sum_{k=1}^{+\infty} |h_{bk}|^2 \left(\lambda_k^2 + R^{-2}\right). \tag{242}$$

Here, C_b is a constant depending on the topography only. Thus we have $E_{mf} \sim E_{fluct}^2 C_b / (Z_2 - Z_b)^2$, which leads to:

$$\lim_{E \to 0} \frac{E_{mf}}{E_{fluct}} = 0, \quad \lim_{E \to 0} \frac{E_{mf}}{E} = 0. \tag{243}$$

The limit $\mathbf{E \to 0}$ with $\mathbf{Z_2 - Z_b \sim C_\alpha E^\alpha}$ with $\alpha \geq 0$: In this limit, we have $E_{fluct} \to 0$. One gets from Fig. 5-a that $\widetilde{\beta} \to \infty$. Hence, from Eqs. (237) and (238), we obtain:

$$\begin{cases} E_{mf} & = C_b \widetilde{\beta}^{-2} + o\left(\widetilde{\beta}^{-2}\right) \\ Z_b - \overline{Z}_2 & = 4 C_b \widetilde{\beta}^{-1} + o\left(\widetilde{\beta}^{-1}\right) \end{cases}, \tag{244}$$

where C_b is defined in Eq. (242). From those two equations along with Eq. (239) and using $Z_2 - Z_b \sim C_\alpha E^\alpha$, we can extract:

$$\widetilde{\beta} \underset{E\to 0}{\sim} \frac{C_\alpha E^\alpha + \sqrt{C_\alpha^2 E^{2\alpha} + 20 C_b E}}{2E}. \tag{245}$$

Now, we have to consider different cases for the value of α.

For $\alpha > 1/2$, we have from Eq. (245) that $\widetilde{\beta} \sim \sqrt{5C_b}E^{-1/2}$. Injecting this in Eq. (244), we gets:

$$\lim_{E\to 0} \frac{E_{mf}}{E} = \frac{1}{5} \tag{246}$$

For $\alpha < 1/2$, we have from Eq. (245) that $\widetilde{\beta} \sim C_\alpha E^{\alpha-1}$. Injecting this in Eq. (244), we gets:

$$\lim_{E\to 0} \frac{E_{mf}}{E} = 0 \tag{247}$$

For $\alpha = 1/2$, we have from Eq. (245) that $\widetilde{\beta} \sim C_{1/2}E^{-1/2}\left(1 + \sqrt{1 + 20C_b/C_\alpha^2}\right)/2$. Injecting this in Eq. (244), we gets:

$$\lim_{E\to 0} \frac{E_{mf}}{E} = \frac{2C_b/C_\alpha^2}{\left(1 + \sqrt{1 + 20C_b/C_\alpha^2}\right)^2}. \tag{248}$$

Contrary to the previous cases, here, the partition of the energy depends on the bottom topography.

Acknowledgements The authors warmly thank M. Potters and A. Licari for their preliminary work on this subject during a traineeship under the supervision of FB. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 616811) (FB and AV). We warmly thank the three referees for their very positive evaluation of our work and for the numerous remark that helped us to improve our work. The level of commitment of the three referee in reading the detail of our computations has been extremely high, and we thank them for this important work.

References

1. Bouchet, F.: Simpler variational problems for statistical equilibria of the 2d euler equation and other systems with long range interactions. Physica D: Nonlinear Phenomena **237**(14), 1976–1981 (2008)
2. Bouchet, F.: Statistical mechanics for geophysical flows. PHD thesis, Université Joseph Fourier-Grenoble (December 2008)
3. Bouchet, F., Sommeria, J.: Emergence of intense jets and Jupiter's Great Red Spot as maximum-entropy structures. Journal of Fluid Mechanics **464**, 165–207 (2002). DOI 10.1017/S0022112002008789
4. Bouchet, F., Venaille, A.: Statistical mechanics of two-dimensional and geophysical flows. Physics Reports **515**, 227–295 (2012). DOI 10.1016/j.physrep.2012.02.001
5. Bourgain, J.: Periodic nonlinear schrödinger equation and invariant measures. Communications in Mathematical Physics **166**(1), 1–26 (1994)
6. Chavanis, P.H., Sommeria, J.: Classification of self-organized vortices in two-dimensional turbulence: the case of a bounded domain. J. Fluid Mech. **314**, 267–297 (1996)
7. Chavanis, P.H., Sommeria, J.: Statistical mechanics of the shallow water system. Phys. Rev. E **65**(2), 026302 (2002)
8. Dubinkina, S., Frank, J.: Statistical relevance of vorticity conservation in the Hamiltonian particle-mesh method. Journal of Computational Physics **229**, 2634–2648 (2010). DOI 10.1016/j.jcp.2009.12.012
9. Dubinkina, S., Frank, J.: Statistical relevance of vorticity conservation in the hamiltonian particle-mesh method. Journal of computational Physics **229**(7), 2634–2648 (2010)
10. Eyink, G.L., Sreenivasan, K.R.: Onsager and the theory of hydrodynamic turbulence. Rev. Mod. Phys. **78**, 87–135 (2006). DOI 10.1103/RevModPhys.78.87
11. Farge, M., Sadourny, R.: Wave-vortex dynamics in rotating shallow water. Journal of Fluid Mechanics **206**, 433–462 (1989)
12. Fofonoff, N.P.: Steady flow in a frictionless homogeneous ocean. J. Mar. Res. **13**, 254–262 (1954)
13. Frank, J., Gottwald, G., Reich, S.: A hamiltonian particle-mesh method for the rotating shallow-water equations. In: Meshfree methods for partial differential equations, pp. 131–142. Springer (2003)
14. Gertz, A., Straub, D.N.: Near-inertial oscillations and the damping of midlatitude gyres: A modeling study. Journal of Physical Oceanography **39**(9), 2338–2350 (2009)
15. Herbert, C.: Additional invariants and statistical equilibria for the 2d euler equations on a spherical domain. Journal of Statistical Physics pp. 1–31 (2013)

16. Herbert, C., Pouquet, A., Marino, R.: Restricted equilibrium and the energy cascade in rotating and stratified flows. Journal of Fluid Mechanics **758**, 374–406 (2014)
17. Majda, A.J., Wang, X.: Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows. Cambridge University Press (2006)
18. Majda, A.J., Wang, X.: The emergence of large-scale coherent structure under small-scale random bombardments. Comm. Pure App. Maths **59**(4), 467–500 (2006)
19. Merryfield, W.J.: Effects of stratification on quasi-geostrophic inviscid equilibria. Journal of Fluid Mechanics **354**, 345–356 (1998)
20. Merryfield, W.J., Cummins, P.F., Holloway, G.: Equilibrium Statistical Mechanics of Barotropic Flow over Finite Topography. Journal of Physical Oceanography **31**, 1880–1890 (2001)
21. Michel, J., Robert, R.: Large deviations for young measures and statistical mechanics of infinite dimensional dynamical systems with conservation law. Communications in Mathematical Physics **159**, 195–215 (1994)
22. Miller, J.: Statistical mechanics of euler equations in two dimensions. Phys. Rev. Lett. **65**(17), 2137–2140 (1990). DOI 10.1103/PhysRevLett.65.2137
23. Miller, J., Weichman, P.B., Cross, M.C.: Statistical mechanics, Euler's equation, and Jupiter's Red Spot. Phys Rev A **45**, 2328–2359 (1992). DOI 10.1103/PhysRevA.45.2328
24. Miranda, A., Barnier, B., Dewar, W.K.: On the dynamics of the zapiola anticyclone. Journal of Geophysical Research: Oceans (1978–2012) **104**(C9), 21,137–21,149 (1999)
25. Nageswaran, V., Turkington, B.: Minmax variational principle for steady balanced solutions of the rotating shallow water equations. Communications in Mathematical Sciences **8**(2), 321–339 (2010)
26. Naso, A., Chavanis, P.H., Dubrulle, B.: Statistical mechanics of two-dimensional Euler flows and minimum enstrophy states. European Physical Journal B **77**, 187–212 (2010)
27. Naso, A., Monchaux, R., Chavanis, P., Dubrulle, B.: Statistical mechanics of beltrami flows in axisymmetric geometry: Theory reexamined. Phys. Rev. E (2010)
28. Onsager, L.: Statistical hydrodynamics. Nuovo Cimento **6 (No. 2 (Suppl.))**, 249–286 (1949)
29. Pedlosky, J.: Geophysical fluid dynamics (1982)
30. Robert, R.: Etats d'équilibre statistique pour l'écoulement bidimensionnel d'un fluide parfait. C. R. Acad. Sci. **1**, 311:575–578 (1990)
31. Robert, R.: A maximum-entropy principle for two-dimensional perfect fluid dynamics. J. Stat. Phys. **65**, 531–553 (1991)
32. Robert, R.: On the statistical mechanics of 2d euler equation. Communications in Mathematical Physics **212**, 245–256 (2000)
33. Robert, R., Sommeria, J.: Statistical equilibrium states for two-dimensional flows. J. Fluid Mech. **229**, 291–310 (1991). DOI 10.1017/S0022112091003038
34. Robert, R., Sommeria, J.: Relaxation towards a statistical equilibrium state in two-dimensional perfect fluid dynamics. Phys. Rev. Lett. **69**(19), 2776–2779 (1992). DOI 10.1103/PhysRevLett.69.2776
35. Salmon, R.: Lectures on Geophysical Fluid Dynamics. Oxford University Press (1998)
36. Salmon, R.: The shape of the main thermocline, revisited. Journal of Marine Research **68**(3-4), 541–568 (2010)
37. Sansón, L.Z., González-Villanueva, A., Flores, L.: Evolution and decay of a rotating flow over random topography. Journal of Fluid Mechanics **642**, 159–180 (2010)
38. Saunders, P.M., King, B.A.: Bottom currents derived from a shipborne adcp on woce cruise a11 in the south atlantic. Journal of Physical Oceanography **25**(3), 329–347 (1995)
39. Smith, W.H.F., Sandwell, D.T.: Global seafloor topography from satellite altimetry and ship depth soundings. Science, p. 1957-1962 **277**, 1957–1962 (1997)
40. Thalabard, S.: Contributions to the statistical mechanics of ideal two and a half dimensional flows (phd thesis). Ph.D. thesis, Université Paris Sud-Paris XI (2013)
41. Thalabard, S., Dubrulle, B., Bouchet, F.: Statistical mechanics of the 3d axisymmetric euler equations in a taylor–couette geometry. Journal of Statistical Mechanics: Theory and Experiment **2014**, P01,005 (2014)
42. Thalabard, S., Saint-Michel, B., E., H., Daviaud, F., Dubrulle, B.: Ferro-turbulence : a statistical mechanics framework for the large-scale structure of turbulent von karman flows. submitted to New Journal of Physics (2014)
43. Touchette, H.: The large deviation approach to statistical mechanics. Physics Reports **478**(1), 1–69 (2009)
44. Turkington, B., Majda, A., Haven, K., Dibattista, M.: Statistical equilibrium predictions of jets and spots on Jupiter. PNAS **98**, 12,346–12,350 (2001)
45. Vallis, G.K.: Mechanisms and parameterizations of geostrophic adjustment and a variational approach to balanced flow. Journal of Atmospheric Sciences **49**, 1144–1160 (1992)
46. Vallis, G.K.: Atmospheric and Oceanic Fluid Dynamics (2006). DOI 10.2277/0521849691
47. Vanneste, J.: Balance and Spontaneous Wave Generation in Geophysical Flows. Annual Review of Fluid Mechanics **45**, 147–172 (2013)
48. Venaille, A.: Bottom-trapped currents as statistical equilibrium states above topographic anomalies. Journal of Fluid Mechanics **699**, 500–510 (2012)
49. Venaille, A., Bouchet, F.: Statistical ensemble inequivalence and bicritical points for two-dimensional flows and geophysical flows. Physical review letters **102**(10), 104,501 (2009)
50. Venaille, A., Bouchet, F.: Oceanic Rings and Jets as Statistical Equilibrium States. Journal of Physical Oceanography **41**, 1860–1873 (2011). DOI 10.1175/2011JPO4583.1

51. Venaille, A., Bouchet, F.: Solvable phase diagrams and ensemble inequivalence for two-dimensional and geophysical turbulent flows. Journal of Statistical Physics **143**(2), 346–380 (2011)
52. Venaille, A., Vallis, G., Griffies, S.: The catalytic role of the beta effect in barotropization processes. Journal of Fluid Mechanics **709**, 490–515 (2012)
53. Warn, T.: Statistical mechanical equilibria of the shallow water equations. Tellus Series A **38**, 1–11 (1986)
54. Weatherly, G.L.: On deep-current and hydrographic observations from a mudwave region and elsewhere in the argentine basin. Deep Sea Research Part II: Topical Studies in Oceanography **40**(4), 939–961 (1993)
55. Weichman, P.B.: Equilibrium theory of coherent vortex and zonal jet formation in a system of nonlinear Rossby waves. Physical Review E **73**(3), 036,313 (2006)
56. Weichman, P.B., Petrich, D.M.: Statistical Equilibrium Solutions of the Shallow Water Equations. Physical Review Letters **86**, 1761–1764 (2001)
57. Xie, J.H., Vanneste, J.: A generalised-lagrangian-mean model of the interactions between near-inertial waves and mean flow. arXiv preprint arXiv:1411.3748 (2014)
58. Young, W., Ben Jelloul, M.: Propagation of near-inertial oscillations through a geostrophic flow. Journal of marine research **55**(4), 735–766 (1997)